SPEECH SIGNAL MODELLING BY A SUM OF COMPLEX FM SIGNALS

A Thesis Submitted
in Partial Fulfilment of the Requirements
for the Degree of
MASTER OF TECHNOLOGY

by

Lt. SANJAY SHARMA

to the

DEPARTMENT OF ELECTRICAL ENGINEERING
INDIAN INSTITUTE OF TECHNOLOGY KANPUR
February 1994

CERFICATE

It is certified that the work contained in the thesis entitled SPEECH SIGNAL MODELLING BY A SUM OF COMPLEX FM SIGNALS by Lt SANJAY SHARMA, has been carried out under my supervision, and that this work has not been submitted elsewhere for the award of a degree.

Pradip Sircar

(Dr P Sircar)
Assistant Professor
Department of Electrical Engineering
Indian Institute of Technology
Kanpur 208 016
India

February 1994

EE-1994-M-SHA-SPE

ABSTRACT

In this thesis, a new model is suggested for modelling signal data from a deterministic or stochastic process. The signal data is modelled as sum of several complex Frequency Modulated Signal. Autoregressive model based on Burg algorithm is used for estimation of model coefficients. These coefficients are used to find the power spectral density. From this the FM subsignals, Carrier and Modulating Frequencies are estimated. Use of Discrete Fourier Transform is made of in estimating the subsignals Modulation Index. Separation of individual subsignal before estimation of Modulation Index is also discussed. These estimated parameters are then used in estimation of subsignals amplitude and Phase.

The model is first fitted on a synthesised data. The model is then fitted on speech signal with varying degree of success.

ACKNOWLEDGEMENT

I would like to express my deepest sense of gratitude and appreciation to Dr P Sircar for his immense support and guidance throughout the span of the work. It was indeed a rewarding experience to carry out research work under his guidance.

I would also like to thank all my colleagues for their support.

Lt. Sanjay Sharma

CONTENTS

LIST OF FIGURES

CHAPTER-1

INTRODUCTION

Parametric modellling of signal is an important step in many signal processing applications. Many practical signals, like image and speech signals , are highly redundant and it may be desirable to represent the signal with a fewer number of samples for economy of storage and /or transmission bandwidth limitations. Knowing the parameters of a signal the resulting parameters can be used as a feature vector for pattern recognition ,speech recognition and alanguage identification etc.

A large number of signals show predominant Frequency Modulation in the time domain.As such they can be assumed as consisting of a combination of several complex Frequency Modulated (FM) singnals, eg speech signals etc. In this thesis an attempt is made to decompose a signal as a combination of various frequency modulated signals.

1.1 NON- STATIONARY SIGNALS

The term stochastic process or random process is used to describe the time evolution of a statistical phenomenon according to probabilistic laws. One particular realization of a discrete time stochastic process is called a time series.For example , the sequence x (n), x (n-1),.....x(n-M+1) represents a time series that consists of present observation x(n) and M-1 past

observation of the process made at time n-1,....., n-M+1 .We say that a stochastic process is strictly stationary if its statistical porperties are invariant to a shift of time , ie the joint probability density function of these observations made at times n, n-1, ..., n-M+1 must remain the same no matter what values we assign to n for fixed M.

In practice, it is not possible to determine the joint probability density function for an arbitrary set of observation made on a stochastic process.As the assumtion of strict stationarity is too restrictive, a notioin of wide sense satıonarity (W.S.S) , is used demanding only the invariance of the first two moments with time.A discreate time stochastic process is said to be wide sense stationary or weakly stationary if its mean is constant and its auto - correlation depends only on the lag ie.

$E[x(n)] = \mu$ and $E[x(n)\, x^{*}(n+k)] = r(k)$; $k = 0, \pm 1, \pm 2..$

Where E = Expectation operator

* = Complex conjugate operator

Any signal which is not wide sense stationary is considered as a non-stationary signal [1] .

1.3 PARAMETER ESTIMATION OF NON-STATIONARY SIGNALS

Parametric method for analysis and synthesis of

stationary signals is a well developed field .Also well developed are Nonparametric methods of nonstationary spectrum estimation. A lot of work has also been done in modelling non-stationary signals within varying degree of success.Adaptive filters have been extensively used for modelling non-stationarity, processes. However, adaptive filters are useful, only if the measure of non-stationarity, which may be the rate of variation of autocovariance function with respect to time is less than the speed of adaptation [2] . If this condition is satisfied , then the error in estimating the weight vector has two components.The first one,called the weight vector noise, is proportional to the speed of adaptation, and the second component called the weight vector lag is inversely proportional to the speed of adaptation. Hence, one optimum operating point is chosen to keep the total error minimum. The conflicting requirements of the components of error , limits its use in the case of highly non-stationary process.

Many deterministic and stochastic discrete time processes encountered in practice are well approximated by a Rational Transfer Function (RTF) model. The RTF model converts the signal representation problem into an equivalent parameter estimation problem Autoregressive (AR), Moving Average (MA) and Autoregressive Moving Average (ARMA) models have been successfully used for stationary signals (4) . Application of these methods to non-stationary signal requires initial processing of signal so as to have a stationary signal . Further , peak of the PSD though occurs at the signal frequencies, but its magnitude is not

proportinal tothe power of the sinusoid as is the case for a Fourier based spectral estimators.The thesis lay out is as described in the following paragraphs.

Chapter 2 deals with the method for frequency determination for the non-stationary model under consideration using the PSD approach. It also discusses the use of Discrete Fourier Transform, for estimation of parameters influencing the power of the signal. A method for resolving overlapped signals in a low point DFT is also discussed.

Chapter 3 deals with the application of the theory developed to synthetic data in cases where the carrier frequencies are wide apart and where they are close.

Chapter 4 deals with the application of the model to speech signals with varying degree of success.

CHAPTER - 2

2.1 FREQUENCY MODULATION

Modulation may be generally defined as the process whereby a message is transformed from its original form, into a signal that is more suitable for transmission and processing to meet needs imposed by specific circumstances. Mathematically, modulation may be described as the process of mapping from a message space to a signal space.

The realisation of the sensitivity of amplitude modulation to noise, interference and distortion lead to the conception of Frequency Modulation (FM) . Frequency modulation is a process in which an information bearing of signal , g(t) is used to vary the instantaneous frequency of another signal c(t). The signal g(t) is referred to as the modulating signal and the signal c(t) as the carrier signal. FM belongs to the class of non-linear modulation unlike AM which belongs to class of linear modulation [6].

Consider frequency modulation of a complex exponential carrier by a complex exponential modulating signal, so that

$$c(t) = A_c e^{j(\Omega_c t + \phi_c)} \text{ and } g(t) = A_m e^{j \Omega_c t}$$

The frequency modulated signal x(t) is given by ,

$$= c(t)\ e^{jk \int g(t)\, dt}$$

$$x(t) = A_c\ e^{j\ (\Omega_c t + \phi c)}\ .\ e^{j\ k \int A_m\ e^{j\Omega_m t}\, dt}$$

$$= A_c\ e^{j\ (\Omega c t + \phi c)}\ .\ e^{\beta\ e^{j\Omega_m t}} \qquad (2\text{-}1)$$

where $\beta = \frac{K\ Am}{\Omega_m} = \frac{\Delta F}{F_m}$ = is the modulation index . More precisely x(t) can written as

$$x(t) = A_c\ e^{\beta\ \cos\ \Omega_m t}\ e^{j\ (\Omega_c t + \phi_c + \beta \sin\ \Omega_m t)} \qquad (2\text{-}2)$$

This equation clearly shows that the modulating signal not only affects the instantaneous frequency of the carrier signal but also its amplitude. That is to say Eqn (2.2) represents an Am-FM combine case.

For only FM consideration we modify the Eqn (2.2) in the following manner

$$g(t) = A_m\ \cos\ \Omega_m t$$

Then

$$x(t) = A_c\ e^{j(\Omega_m t + \phi_c)}\ e^{j\beta\ \sin\ \Omega_m t}$$

$$= A_c\ e^{j\ (\Omega_c t + \phi_c + \beta\ \sin\ \Omega_m t)} \qquad (2\text{-}3$$

A discrete time version of Eqn (2.3) will be given by

$$x[n] = A_c\ e^{j(\omega_c n + \phi_c + \beta\ \sin\ \omega_m n)};$$

$$n = 0, 1, \ldots, N-1 \qquad (2\text{-}4$$

Where $\omega = \Omega T$ is the discrete frequency with T as the sampling interval.

2.2 SIGNAL MODELLING

We will consider here a signal model comprising of a superposition of several frequency modulated subsignals. A complex signal x(t) consist-ing of P complex exponential carriers with sinusoidal modulation can be expressed as

$$x(t) = \sum_{i=1}^{P} A_i \, e^{j\,(\Omega_i t + \phi_i + \beta_i \sin N_i t)} \qquad (2\text{-}2.1)$$

where A_i = amplitude of i^{th} exponential carrier

ϕ_i = phase of i^{th} exponential carrier, which is a random variable uniformly distributed $(0, 2\pi)$.

β_i = modulation index of i^{th} exponential

Ω_i = frequency of i^{th} exponential carrier

Ni = frequency of i^{th} baseband singal

Satisfying the sampling theorem, an appropriate sampling frequency with time period T is chosen for the uniform sampling of x(t). The discrete-time samples of the singal are of the form

$$x[n] = \sum_{i=1}^{P} A_i \, e^{j\,(\omega_i n + \phi_i + \beta_i \sin \nu_i n)} \qquad (2.2.2)$$

where w_i = i^{th} discrete angular carrier frequency

ν_i = i^{th} discrete angular modulating frequency

ϕ_i = phase of i^{th} exponential which is a r.v uniformly distributed (o , 2π).

The purpose of the thesis is to perform anlysis of a given complex signal so that it can be expressed as a sum of several FM subsignals with complex exponential carriers and sinusoidal baseband singals. In otherwards, the parametric modelling of the signal x[n] will be achieved by determining the parameters of equation (2.2.2). The steps involved will be :

1 . Estimation of the carrier and modulating frequencies .

2 . Estimation of the modulation index of each sub-signal .

3 . Estimation of the amplitude and phase of each sub-signal .

2.3 MODIFIED AUTOCORRELATION FUNCTION

For the signal x[n] which is a sum of complex FM signals, the autocorrelation function is given by

$$r_x[n,k] = E[\, x^*[n]\ x[n+k]\,] \qquad (2.3.1)$$

where k is the lag

from Eqn (2.2.2) we get

$$x^*[n] = \sum_{i=1}^{p} A_i\, e^{-j(\omega_i n + \phi_i + \beta_i \sin \nu_i n)} \qquad (2.3.2)$$

$$x[n+k] = \sum_{i=1}^{p} A_i\, e^{j(\omega_i (n+k) + \phi_i + \beta_i \sin(\nu_i (n+k)))} \qquad (2.3.3)$$

from Eqn (2.3.1), (2.3.2), and (2.3.3) we have

$$r_x[n,k] = E\left[\sum_{i=1}^{p} A_i\, e^{-j(\omega_i n + \phi_i + \beta_i \sin \nu_i n)}\right.$$

$$\left.\sum_{l=1}^{p} A_l\, e^{j\,(\omega_l(n+k)\,\phi_l + \beta_l \sin \nu_l n)}\right] \qquad (2.3.4)$$

$$E\left[e^{j\,(\phi_l - \phi_i)}\right]_{(\phi_l \neq \phi_i)} = \frac{1}{2\pi}\int_0^{2\pi} e^{j\,(\phi_l - \phi_i)}\, d\,(\phi_l - \phi_i)$$

$$= 0 \qquad (2.3.5)$$

$$E\left[e^{j\,(\phi_l - \phi_i)}\right]_{(\phi_l = \phi_i)} = E[1] = 1 \qquad (2.3.6)$$

In view of Eqns (2.3.5) and (2.3.6) , Eqn (2.3.4) becomes

$$r_x[n,k] = \sum_{i=1}^{p} A_i^2\, e^{jw_l k}\; e^{j2\beta_i \sin(\nu_i \frac{K}{2})} \cos\left(\nu_l \frac{(2n+k)}{2}\right)$$

(2.3.7)

Because of its double dependence on position n and lag k, the autocorrelation function, $r_x[n,k]$ can not be used in determination of the frequencies (carrier and modulating) of the signal. Also, since discrete values of x[n] are available for only one observation, ensemble average can not be taken.

To overcome these shortcomings of the time-dependent autocorrelation function, we take the product function

$P_x[k]$, defined as

$$P_x[k] = \left[x^*[n]\;\; x[n+k]\right]_{n=-\frac{-k}{2}}$$

$$k = 0, \pm 2, \pm 4, \pm 6, \ldots \qquad (2.3.8)$$

$$P_x[k] = \left[\sum_{i=1}^{p} A_i\, e^{-j\,(\omega_l n + \phi_l + \beta_i \sin \nu_l n)}\right.$$

$$\sum_{l=1}^{p} A_l \, e^{j(\omega_l(n+k) + \phi_l + \beta_l \sin \nu_l (n+k)]} \qquad (2.3.9$$

$$= \sum_{i=1}^{p} \sum_{l=1}^{p} A_i A_l \, e^{j(\omega_l n - \omega_i n + \omega_l k)} \, e^{j(\phi_l - \phi_i)} \cdot$$

$$e^{j(\beta_l \sin(\omega_l(n+k)) - \beta_i \sin(\nu_i+n))} \qquad (2.3.10)$$

To have further insight of Eqn (2.3.10), we deal with the specific case of say p=2. then

$$P_x[k] = A_1^2 \, e^{j(\omega_1 k + 2\beta_1 \sin \frac{\nu_1 k}{2})}$$

$$+ A_2^2 \, e^{j(w_2 k + 2\beta_2 \sin \frac{\nu_2 k}{2})}$$

$$+A_1A_2e^{j[\frac{(\omega_1+\omega_2)k}{2} + (\phi_2 - \phi_1) + \beta_1 \sin(\nu_1 - \frac{k}{2}) + \beta_2 \sin(\nu_2 - \frac{k}{2}}$$

$$+A_1A_2e^{j[\frac{(\omega_1+\omega_2)k}{2} + (\phi_1 - \phi_2) + \beta_1 \sin(\nu_1 - \frac{k}{2}) + \beta_2 \sin(\nu_2 - \frac{k}{2}}$$

(2.3.11)

Let us a alyse first term of Eqn (2.3.11)

$$A_1^2 \, e^{j(\omega_1 k + 2\beta_1 \sin \frac{\nu_1 k}{2})} \qquad (2.3.11a)$$

Using the properties of bessel functions we know [2].

$$e^{jx\sin\phi} = \sum_{l=-\infty}^{\infty} J_l(x)\, e^{jl\phi} \qquad (2.3.12)$$

Where coefficients $J_l(x)$ are Bessel functions of the first kind, of order l and argument x. Hence Eqn (2.3.11a) can be rewritten as

$$A_1^2\, e^{j(\omega_1 k)} \sum_{l=-\infty}^{\infty} J_l(2\beta)\, e^{jl\frac{\nu_1 k}{2}} \qquad (2.3.13)$$

$$A_1^2 \sum_{=-\infty} J_l(2\beta_1)\, e^{j(w_1 k + l\frac{\nu_1 k}{2})},\ k=0, \pm 2, \pm 3,... \qquad (2.3.13$$

Similarly the other terms of Eqn (2.3.11) can be expessed in terms of the Bessel function, giving the expression of $P_x[k]$ i.e. Eqn (2.3.11)

$$P_x[k] = A_1^2 \sum_{l1=-\infty}^{\infty} J_{l1}(2\beta_1)\, e^{j(\omega_1 k + \frac{l_1 \nu_1 k}{2})}$$

$$+ A_2^2 \sum_{l2=-\infty}^{\infty} J_{l2}(2\beta_2)\, e^{j(w_2 k + \frac{l_2 \nu_2 k}{2})}$$

$$+ A_1 A_2\, e^{j(\phi_2-\phi_1)} \sum_{l1=-\infty}^{\infty} \sum_{l2=-\infty}^{\infty} J_{l1}(2\beta_1)\, J_{l2}(2\beta_2)\, e^{j((\frac{\omega_1+\omega_2}{2})k + (\frac{l_1\nu_1+l_2\nu_2}{2})k)}$$

$$+ A_2 A_1\, e^{j(\phi_1-\phi_2)} \sum_{l1=-\infty}^{\infty} \sum_{l2=-\infty}^{\infty} J_{l1}(2\beta_1) J_{l2}(2\beta_2)\, e^{j((\frac{\omega_1+\omega_2}{2})k + (\frac{l_1\nu_1+l_2\nu_2}{2})k)} \qquad (2.3.14)$$

The following observations can be made from Eqn (2.3.14)

(1) The spectrum of P_x [k] will contain the carrier components occuring at twice the original values as k takes only even values and an infinite set of true modulating frequencies located symmetrically on either side of the carriers at frequency separation of ν_m, $2\nu_m$, $3\nu_m$....

(2) The amplitude of the various component varies with modulation index.

(3) Spectrum will also contain cross carriers which are the mean of carriers and cross modulating frequencies .

The magnitude of Bessel functions $J_n(\beta)$ become very small for $|n| > |\beta|$, [6]. For all practical purposes $|n| \cong |\beta+1|$, is the largest n which will have any appreciable value for $J_n(\beta)$.[6] . Therefore in the summation of Eqn (2.3.14) for practical purposes only those terms for (n) extending from zero to little greater than β need be included.

2.4 ESTIMATION OF CARRIER AND MODULATING FREQUENCYES

2.4.1 POWER SPECTRAL DENSITY TECHNIQUES

The power Spectial Density (PSD) of a complex WSS random process x[n] is defined as

$$P_{xx}(f) = \sum_{k=-\infty}^{\infty} r_{xx}(k) \exp(-j2\pi fk); \qquad -\frac{1}{2} \le f \le \frac{1}{2} \qquad (2.4.$$

The frequency (f) may either be thought of as the fraction of the sampling frequency used in obtaining the data samples from a continuous process or as the number of cycles/sample. The PSD

function describes the distribution of power with frequency of the process under consideration. Since the PSD is a function of an infinite number of ACF values, the attempt of estimating the PSD based on a finite data set is erroneous. At best we can only estimate a limited number of ACF values, and therefore, parametric modelling is employed to extrapolate the ACF sequence for all lags.

This allows us to replace spectral estimation problem by a parameter estimation problem. It is critical, however that the model be an accurate representation of the underlying process, at least as far as the PSD is concerned.

Thus the spectral estimation, in the context of modelling becomes a three step procedure. The first step is to select a model. The second step is to estimate the parameters of the assumed model using the available data samples. The third step is to obtain the spectral estimate by substituting the estimated model parameters into the theoretical PSD implied by the model [3].

Many deterministic and stochastic discrete time processes encountered in practise are well approximated by a rational transfer function model. In this model input driving sequence u[n] and the output sequence x[n] that is to model the data are related by the linear difference equation

$$x[n] = -\sum_{k=1}^{p} a[k]\, x[n-k] + \sum_{k=0}^{q} b[k]\, u[n-k] \qquad (2.4.2.2)$$

The most general linear model is termed as an ARMA model. The transfer function H(z) between the input u[n] and the output x[n] for the ARMA process is the rational function

$$H(z) = \frac{B(z)}{A(z)} = \frac{\sum_{k=0}^{q} b[k]\, z^{-k}}{\sum_{k=0}^{p} a[k]\, z^{-k}}$$

Often the driving process is assumed to be a white noise sequence of zero mean and variance σ. The PSD of the ARMA output process becomes

$$P_{ARMA}(f) = \sigma^2 \left| \frac{B(f)}{A(f)} \right|^2$$

where $A(f) = A(z)\,|_{z=e^{j2\pi f}}$

$B(f) = B(z)\,|_{z=e^{j2\pi f}}$

If all the a[k] coefficients, except a[0]=1, vanish for ARMA parameters, then the process is strictly a MA process of order q and

$$P_{MA}(f) = \sigma^2 |B(f)|^2$$

If all the b[k] coefficients except b[0]=1, vanish for ARMA parameters the process is strictly an AR process of order p and

$$P_{AR}(f) = \frac{\sigma^2}{|A(f)|^2}$$

The AR model is appropriate for spectra with sharp peaks but not deep valleys. The MA model can be used for spectra with deep valleys but not sharp peaks. The complex FM signal model spectra

will yield many pairs of carriers and carrier plus modulating frequencies. The AR model therefore appears to be the most appropriate for this model.

The selection of model order in AR spectral estimation is a critical one. Too low an order results in a smoohed estimate, while too large an order causes suprious peaks and general statistical instablity.

For the complex FM signal model the product function (P[k]; k=0, ± 2, ± 4, ...) are fitted onto an AR model based on Burg algorithm [3]. The AR parameters so obtained are used to compute the PSD of the product function from Eqn (2.3.14). The PSD should consist of several sets of peaks, each set consisting of a single carrier freqency which is twice the carrier frequency followed by true modulating frequencies symmetrically , distributed about the carrier with diminishing magnitudes. Also present will be peaks representing cross term carrier which are sum of the true carrier frequencies followed by the cross modulating frequencies. If the data is real then the PSD will be symmetrical about frequency f=0 and accordingly the AR model order chosen should also be twice the number of frequency pairs expected.

2.5 ESTIMATION OF MODULATION INDEX

From Eqn (2.3.14) it is clear that the modulation index β, for a given set of carrier and modulating frequencies and lag will affect the magnitude of the PSD at those frequencies. The magnitude however, is not proportional to the power of the

sinusoid as in the case for a fourier based spectral estimator but to the square of power. For a single complex sinusoid in noise with SNR ϑ and model order p, in the limiting case for $\vartheta np >> 1$ and $p >> 1$

$$P_{AR}(f)_{max} \alpha \sigma_w^2 (\vartheta p)^2 \qquad (2.5.1)$$

The minimum variance spectral estimate though having a resolution poorer than AR spectral estimate has magnitude directly proportional to power of the sinusoid for a given model order

ie. $P_{MVSE}(f) \alpha A^2$ (2.5.2) [KAY]

Both Eqn (2.5.1) and (2.5.2) clearly depicts that the magnitude of the spectral estimate is dependent on the model order. We utilise the fourier transform approach to evaluate the modoulation index.

Given a complex valued data sequence x[n] of length N, the Discrete Fourier Transform is given by [5]

$$X[K] = \sum_{n=0}^{N-1} x[n] \; e^{-j\frac{2\pi kn}{N}} \quad ; \; 0 \leq K \leq N-1 \qquad (2.5.1$$

Substituting value of x[n] from Eqn (2.2.2)

$$X[K] = \sum_{n=0}^{N-1} \left(\sum_{i=1}^{p} A_i \; e^{j(w_i n + \phi_i + \beta_1 \sin \nu_1 n)} \right) e^{-j\frac{2\pi kn}{N}} \qquad (2.5.2)$$

Let $A_{ci} = A_1 \; e^{j\phi_i}$

From eq (2.3.12), we can write

$$e^{j\beta \sin \nu n} = \sum_{l=-\infty}^{\infty} J_l(\beta)\, e^{jl\nu n}$$

Hence eq (2.5.2) can be written as

$$X[K] = \sum_{n=0}^{N-1} \left(\sum_{i=1}^{p} A_{ci}\, e^{j\,\omega_1 n} \sum_{l_i=-\infty}^{\infty} J_{l_1}(\beta_1)\, e^{jl\nu_1 n} \right) e^{-j\,\frac{2\pi kn}{N}}, \qquad (2.5.3)$$

Rearranging the summation signs

$$= \sum_{i=1}^{p} A_{ci} \sum_{l_1=-\infty}^{\infty} J_{l_i}(\beta_i) \sum_{n=0}^{N-1} e^{j(\omega_i - \frac{2\pi k}{N} + l\nu_1)n} \qquad (2.5.4)$$

$$\text{But} \sum_{n=0}^{N-1} e^{j(w_1 - \frac{2\pi k}{N} + l\nu_1)\, n} = \frac{1-e^{j(\omega_1 - \frac{2\pi k}{N} + l\nu_i)N}}{1-e^{j(w_1 - \frac{2\pi k}{N} + l\nu_i)}}$$

Hence

$$X[K] = \sum_{i=1}^{p} A_{ci} \sum_{l_i-\infty}^{\infty} J_{l_i}(\beta_1) \left[\frac{1-e^{j(w_i - \frac{2\pi k}{N} + l\nu_1)\, N}}{1-e^{j\,(w_i - \frac{2\pi k}{N} + l\nu_1}} \right] \qquad (2.5.5)$$

or

$$|X[K]| = \sum_{i=1}^{p} A_i \sum_{l_i=-\infty}^{\infty} J_{l_1}(\beta_1) \left| \frac{1-e^{j(w_1 - \frac{2\pi k}{N} + l\nu_i)N}}{1-e^{j(w_i - \frac{2\pi k}{N} + l\nu_i}} \right| \qquad (2.5.6)$$

Magnitude of the DFT ie $|X[K]|$ will have peak values at those K_1 where for $l_i = 0, \pm 1, \pm 2....$

$$w_1 - \frac{2\pi k}{N} + l\nu_i = 0$$

$$\text{or } K_i = (w_i + l\nu_1)\, \frac{N}{2\pi}\ ;\ l=0, \pm 1, \pm 2,....$$

$$0 \le K_1 \le N-1 \qquad (2.5.7)$$

Expanding Eqn (2.5.6) as

$$|\ x\ [k]| = A_1 \sum_{l=-\infty}^{\infty} J_1(\beta_1)\, r_{11} + A_2 \sum_{l=-\infty}^{\infty} J_1(\beta_2)\, r_{2,1} + ...$$

$$+ A_p \sum_{l=-\infty}^{\infty} J_1(\beta_p)\, r_{p,1} \qquad (2.5.8)$$

where $r_{1,1} = \left| \frac{1-e^{j(w_1 \frac{2\pi k}{N} + l\omega_i)N}}{1-e^{j(w_i - \frac{2\pi k}{N} + l\omega_1)}} \right|$

At the 1st highest peak corresponding to $l_1 = 0$,

$$| X [k]| = A_1 J_0(\beta_1) k_{1o} + A_2 J_{12}(\beta_2) K_{2,0} + \ldots + A_p J_{1p}(\beta_p) K_{P,0} \quad (2.5.9)$$

Assuming that the values $J_{12}(\beta_2)$, $J_{13}(\beta_3)\ldots$ $J_{1p}(\beta_p)$ are sufficiently small, their contribution towards $| X [k]|$ can be neglected. Hence Eqn. (2.5.9) can be rewritten as

$$| X [k1] | = A_1 J_o (\beta_1) r_{1,0} \quad (2.5.10)$$

similarly at $l = +1$ and -1

$$| X [k2]| = A_1 J_1 (\beta_1) r_{1,1} \text{ and} \quad (2.5.11)$$

$$| X [k2]| = |A_1 \underline{J}_1 (\beta_1) r_{1,-1}|$$

$$= A_1 J_1 (\beta_1) r_{1,-1} \quad (2.5.12)$$

Taking the ratio of Eqn (2.5.10) and (2.5.11) we get

$$\frac{A_1 J_0(\beta_1) r_{1,0}}{A_1 J_1(\beta_1) r_{1,1}} = \left| \frac{X [K1]}{X [k2]} \right|$$

$$\text{or } \frac{J_0(\beta_1)}{J_1(\beta_1)} = \left| \frac{r_{1,1} X [k1]}{r_{1,0} X [k2]} \right| \quad (2.5.13)$$

Similarly we can get $\frac{J_0(\beta_1)}{J_2(\beta_1)}$ and $\frac{J_0(\beta_p)}{J_1(\beta_p)}$

Once the ratio of $J_0(\beta)/J_1(\beta)$ is known we can get the value of β, the modulation index from the tables having values of $J_0(x)/J_1(x)$ for different values of (x).

2.6 ESTIMATION OF AMPLITUDE AND PHASE

Once the true carrier and modulating frequencies and corresponding modulation index are determined by the method discussed above, the remaining probelm of estimating the amplitudes and phases turns out to be a linear estimation problem.

Given a discrete set of observation points of a deterministic or stochastic process { x[n], n = 0,1..., N-1} we can fit them into the discussed AR model by Eqn (2.2.2) as :

$$x[n] = \sum_{i=1}^{p} A_i \, e^{j\,(w_i n + \phi_i + \beta_i \sin \nu_i n)}$$

$$= \sum_{i=1}^{p} A_{ci} \, e^{j(w_i n + \beta_i \sin \nu_i n)}$$

$$= \sum_{i=1}^{p} A_{ci} \, \delta_{n,i} \qquad (2.5.1)$$

where $A_{ci} = A_i \, e^{j\phi_i}$ is the complex carrier amplitude of i^{th} exponential.

$$\delta_{n,i} = e^{j(w_i n + \beta_i \sin \nu_i n)}$$

Eqn (2.6.1) can be written as a matrix equation as follows

$$\begin{bmatrix} x[0] \\ x[1] \\ \vdots \end{bmatrix} = \begin{bmatrix} \delta_{0,1} & \delta_{0,2} \cdots\cdots\cdots \delta_{0,p} \\ \delta_{1,1} & \delta_{1,2} \quad\quad\quad \delta_{1,p} \\ \vdots & \vdots \quad\quad\quad \vdots \end{bmatrix} \begin{bmatrix} A_{c1} \\ A_{c2} \\ \vdots \end{bmatrix}$$

$$\begin{bmatrix} x[N-1] \end{bmatrix}_{NX1} \quad \begin{bmatrix} \delta_{N-1,1} & \delta_{N-1,2} & \delta_{N-1,p} \end{bmatrix}_{NXp} \quad \begin{bmatrix} A_{cp} \end{bmatrix}_{PX1}$$

or $\underline{X} = Z\,\underline{A}$ (2.6.2)

Z is a NXp rectangular matrix with N >> p. To solve eqn (2.6.2) we need to compute the inverse of Z. We solve Eq (2.6.2) by using the pseudo inverse of Z. The singular value Decomposition (SVD) of matrix Z can be written as [7].

$$Z_{N\ X\ p} = Q_{N\ X\ N}\, D_{N\ Xp} V^{H}_{\ p\ X\ p} \qquad (2.6.3)$$

where Q has the left singular vectors of Z,
D is a dia-gonal matrix containing the singular values of Z and V has the Right singular vectors of Z. Q and V are unitary matrices.

$$Z = [\,q_1, q_2 \cdots q_n]_{N\ X\ N} \begin{bmatrix} \sigma_1 & 0 \cdots\cdots 0 \\ 0 & \sigma_2 \quad 0 \\ \vdots & \quad\ddots\quad \vdots \\ 0 & \cdots\ \sigma_p \end{bmatrix}_{NXp} \begin{bmatrix} v_1^H \\ v_2^H \\ \\ \\ v_p^H \end{bmatrix}_{p\ X\ p}$$

where q_i ad v_i are column vectors and σ_i are the signular values of Z with $\sigma_1 \geq \sigma_2 \geq \sigma_n \geq 0$. The psendoinverse of Z is given by

$$Z^{\#} = V_{p\ x\ p}\ D^{-1}_{\ p\ x\ N}\ Q^{H}_{N\ x\ N} \qquad (2.6.4)$$

The solution of Eq (2.6.2 is given as

$$\underline{A} = Z^{\#}\,\underline{X} \qquad (2.6.5)$$

The magnitude of A_{ci} will give us the information about amplitude, and phase of A_{ci} will give the information about phase of the i^{th} exponential carrier.

2.7 CLOSE FREQUENCY CASE

With establishment of the procedure of decomposing a signal as a sum complex exponential FM subsignals, we now proceed to a more complex case wherein the carrier frequencies are so close that the pattern of one overlaps with that of another. Clearly in such a situation, though the carrier and modulating frequencies can be ascertained from the PSD plots as explained before, the

determination of modulation index, β , becomes more difficult as the magnitude of overlapping frequencies get altered from its true value.

Under such circumstances the advantage used in obtaining Eq (2.5.10) from Eq (2.5.9) no longer holds good. As such it is must that pattern of each carrier and set of modulating frequencies on either side should be separated from one another, ie the affect of one waveform over another in DFT plot has to be removed.

To overcome this problem, we bank on the properties of DFT, namely the real part of DFT is of even symmetry and imaginary part is of odd symmetry [5]. So keeping the carrier frequency as centre

we flip about other in forward direction or backward direction, thus creating the pattern about the carrier frequency. This pattern is then deducted from the overall signal thus removing its affect. Same process is carried on the residual signal till all patterns are separated.

The process of flipping is started from one extreme carrier frequency as its frequency will be affected the least because of overlapping. Once the patterns are separated the value of the modulation index can be found by Eqn. (2.5.13).

Due to dependence of resolution of DFT on the number of point of DFT taken, it is possible that at a moderate point DFT, two distinc-t peaks may appear as a single peak in the DFT magnitude plot. This type of possibility is predominantly high in side frequencies of two adjacent patterns. Under such circumstances using the above discussed flipping procedure will introduce a new frequency in the residual signal, which was not in the signal originally.

To overcome this problem, we make an estimate of all possible frequencies whose combination can lead to the observed frequency, in the DFT magnitude plot. This is most easily done by joining the adjacent frequencies by linear interpolation. In short an envelope of the DFT magnitude plot for the frequencies of interest is obtained.

The process of eliminating the affect of overlapping is then done by flipping as discussed above. After each flip, the spurious frequency appearing in the residual signal is removed and the process of flipping carried till all patterns are well resolved.

CHAPTER – 3

SIMULATION RESULTS

3.7 Introducation

To show the effectiveness of the model discussed in chapter 2, we do some simulation study with syntnesised data. A signal consisting of several complex FM sub signals is sampled and the product function for different lags are computed. These are then fitted on an AR model whose coefficients are computed by Burg algorithm. As expected the PSD of the product function consists of sets of frequencies, with a high peak at twice the original carrier frequency with diminishing side peaks at true modulating frequency from the carrier.

3.2 SIGNAL DATA 1

The complex FM signal used for the purpose of simulation consists of two complex FM signals

$$x(t) = \sum_{i=1}^{2} A_i \, e^{j\,(\Omega_i t + \phi_i + \beta_i \sin N_i t)}$$

where

$$A_1 = 1,\ \Omega_1 = 800\,\pi,\ \phi_1 = 30^0,\ \beta_1 = .4,\ N_1 = 70\,\pi$$

$A_2 = 1,\ \Omega_2 = 1600\ \pi,\ \phi_2 = 10^0,\ \beta_2 = 1,\ N_2 = 66\ \pi$

Bandwidth (BW) of FM singal = $2(\beta+1)$ fm

ie $BW_1 = 98$ Hz

$BW_2 = 132$ Hz

Nyquist Rate 1 = 1992 Hz

Nyquist Rate 2 = 3728 Hz

By sampling theorem sampling frequency (Fs) ≥ Nyquist Rate

ie Fs = 4000 Hz

Satisfying the sampling theorem, the sampling frequency correponding to the period T=.25 m sec is chosen for uniform sampling grid. The samples taken are of the form

$$x[n] = \sum_{i=1}^{2} A_1\ e^{j\ (\Omega_1\ Tn + \phi_1 + \beta_i\ \sin N_1 Tn)};\ n = 0,1,\ldots \qquad (3.2.1)$$

The real and imaginary parts of the signal are plotted in Fig 3.2.1 and Fig 3.2.2 respectively.

3.2.1 SIMULATION RESULT

Using the smaples generated by Eqn (3.2.1), the product function $P_x[K]$ were calculated by Eqn (2.3.8) for K = (−100, −98, ... −2, 0, 2, ... 98, 100) a total of 101 points. The real and imaginary parts of $P_x[K]$ are plotted in Figs 3.2.3 and 3.2.4 respectively.

The PSD plot can then be obtained by utilizing the AR

coefficients. In the PSD plot, if the carrier and carrier plus modulating frequencies are not well resolved then we have to increase the model order to split them. However, if the AR model order chosen is too high then line splitting will occur ie. a single frequency may falsely split into two lines giving rise to spurious peaks.

Figure 3.2.5 shows the effect of AR model order on the PSD plot. From the figure, for model order 50, the desired patterns are found centered at discrete frequencies

$2f_{c1} = .2, \qquad 2f_{c2} = .3, \qquad 2f_{c3} = .4$

Hence $f_{c1} = .1, \qquad f_{c2} = .15, \qquad f_{c3} = .2$

Clearly, the pattern at $f_{c2} = .15 = \frac{.1 + .2}{2}$ represents the cross term as discussed in Eqn (2.3.4), and as such of no interest to us. Hence the two carrier frequencies are $f_{c1} = .1, \quad f_{c2} = .2.$

Also from the same figure the side peaks give the modulating frequencies as

$f_{m1} = .014, \qquad f_{m2} = .012$

For evaluation of the modulation index ie β, we use the DFT method discussed in section 2.5.

A 512 point DFT of the signal x[n] is plotted in Fig 3.2.6. Calculating k (sample no:) corresponding to l=0, ±1, ±2, ...

using Eqn (2.5.7) and finding ratio Jo (β)/J1 (β), Jo (β)/J2 (β) etc using Eqn (2.5.13) for both patterns, in a tabular from we have,

Set No	l=-2 K=	l=-1 K=	l=0 K=	l = +1 K =	L = +2 K =	Jo/J1	Jo/J2	β
1.	-	55.68	51.2	46.72	-	4.555		.4
2.	93.952	98.176	102.4	106.624	110.85	1.751	6.219	1

The amplitudes and phase of the two signals are calculated using Eq (2.6.5) which give

$A_1 = 1, \qquad \phi_1 = 30.1^0$

$A_2 = 1, \qquad \phi_2 = 9.99^0$

3.3 SIMULATION NO 2

This simulation with the synthetic data deals with the procedure dealt in section 2.7. The complex FM signal used for the purpose consists of two complex FM signal with parameters as,

$A_1 = 1, \Omega_1 = 1400\,\pi, \phi_1 = 30^0, \beta_1 = .4, N_1 = 70\,\pi$

$A_2 = 1, \Omega_2 = 1600\,\pi, \phi_2 = 10^0, \beta_2 = 1, N_2 = 66\,\pi$

Satisfying the sampling theorem, the sampling frequency corresponding to the period T=.25 m sec is chosen for uniform sampling grid. The samples taken are of the form of Eqn 3.2.1.

Using the samples generated the product function $P_x[K]$ were

calculated by Eqn (2.3.8) by K = (−100, −98, ... −2, 0, 2, ... 98, 100). The product function (P_x [k]) are then fitted in an AR model which uses Burg algorithm. From the PSD shown in figure 3.3.1 the carrier and modulating frequencies are known.

$f_{c1} = .175$ $\quad$ $f_{c2} = .2$

$f_{m1} = .014$ $\quad$ $f_{m2} = .012$

A 1000 point DFT magnitude is plotted in Fig 3.3.2. By comparison with the patterns about the two carriers it is evident that the magnitude of frequencies between the two carriers have altered due to overlapping.

Fig 3.3.3 shows the envelope obtained by linearly interpolating between the peaks. The interpolation is done on both the real and imaginary values of the complex DFT. Then they are mixed to find the magnitude. Working on real and imaginary part of DFT becomes necessary as the magnitude plot does not give any information about the manner of combination of two different complex values.

First backward flip is carried about $f_{c2} = .2$. Pattern about this carrier frequency is well established by the right most curve in the Fig (3.3.4). This pattern is then deducted from the original signal envelope. The residual consists of the well desired pattern about frequency $f_{c1} = .175$, shown by the left curve of Fig (3.3.4). The spurious signal from the residual has been eliminated.

Calculation of modulation index, from Fig (3.3.4) and Eqn (2.5.10) gives

$$\frac{J_0(\beta_1)}{J_1(\beta_1)} = 5.134$$

which from tables give

$$\beta_1 \cong .38$$

and for the other pattern

$$\frac{J_0(\beta_2)}{J_1(\beta_2)} = 1.774 \qquad \frac{J_0(\beta_2)}{J_2(\beta_2)} = 6.61$$

which from tables gives

$$\beta_2 \cong .99$$

The amplitudes and phase of the two signals can be calculated using Eqn (2.6.5)

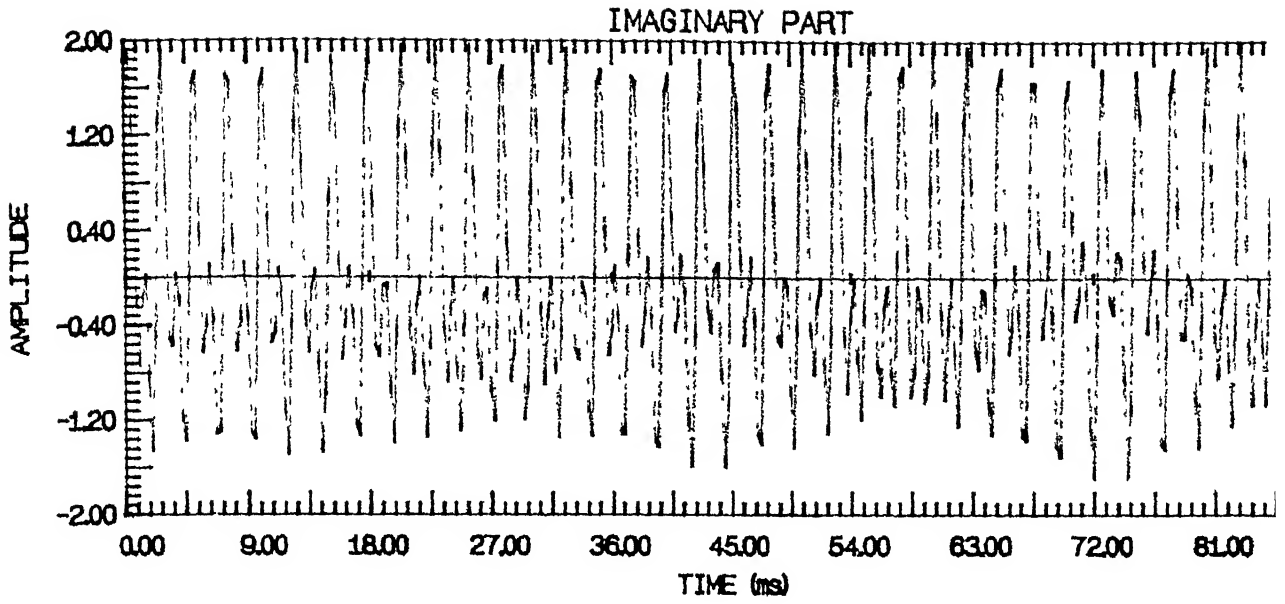


Fig. 3.2.1 Real Part of complex FM signal

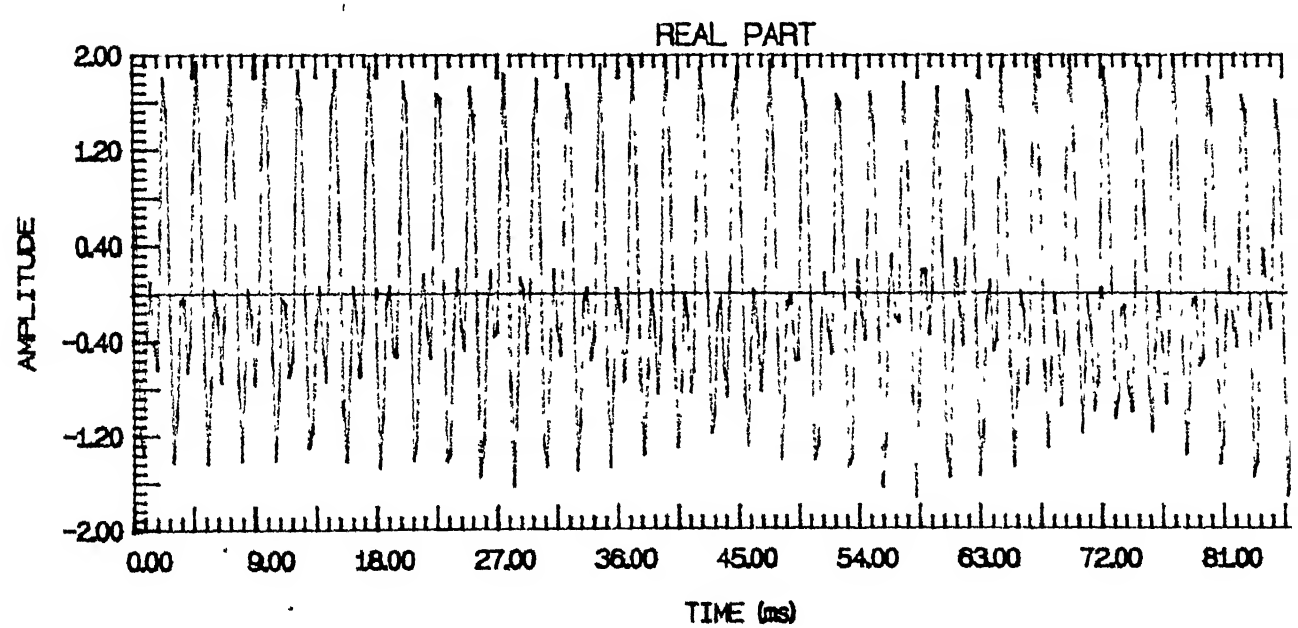


Fig. 3.2.2 Imaginary Part of complex FM signal

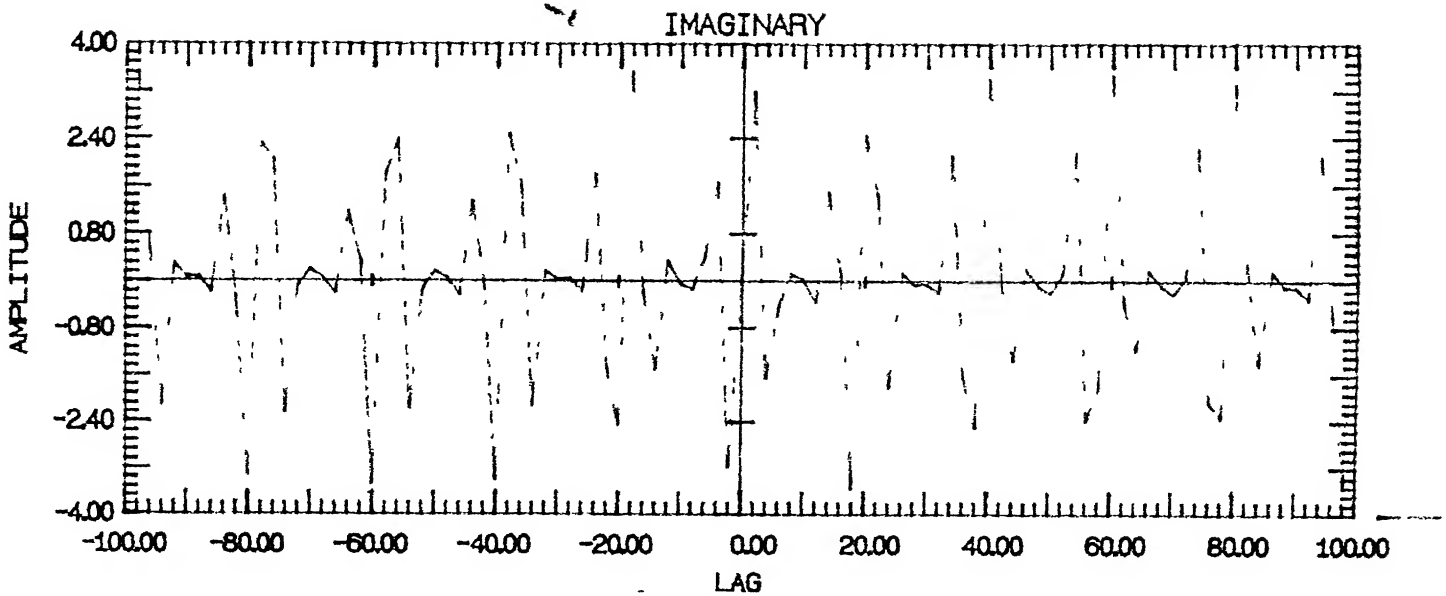


Fig. 3.2.4 Imaginary Part of Product function

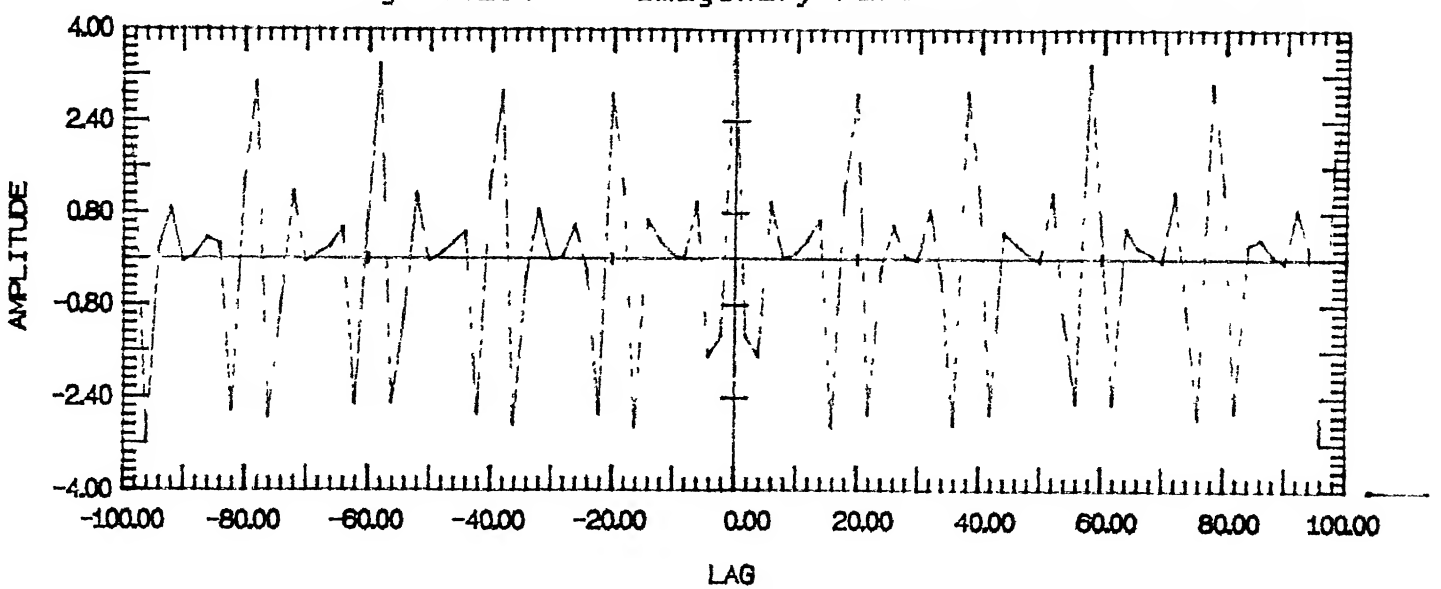


Fig. 3.2.3 Real Part of Product function

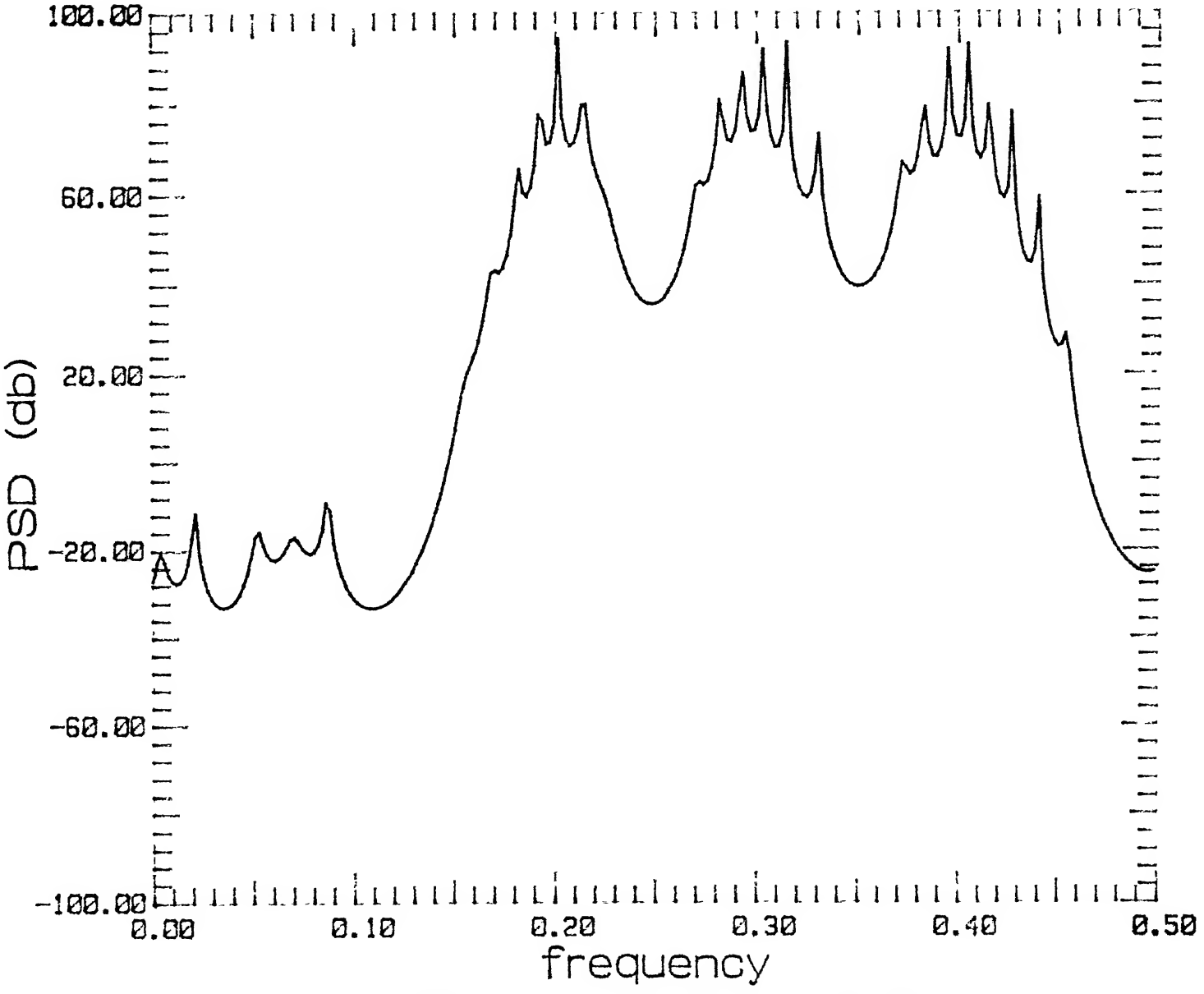


Fig. 3.2.5a PSD : K = ± 100, Model order 50

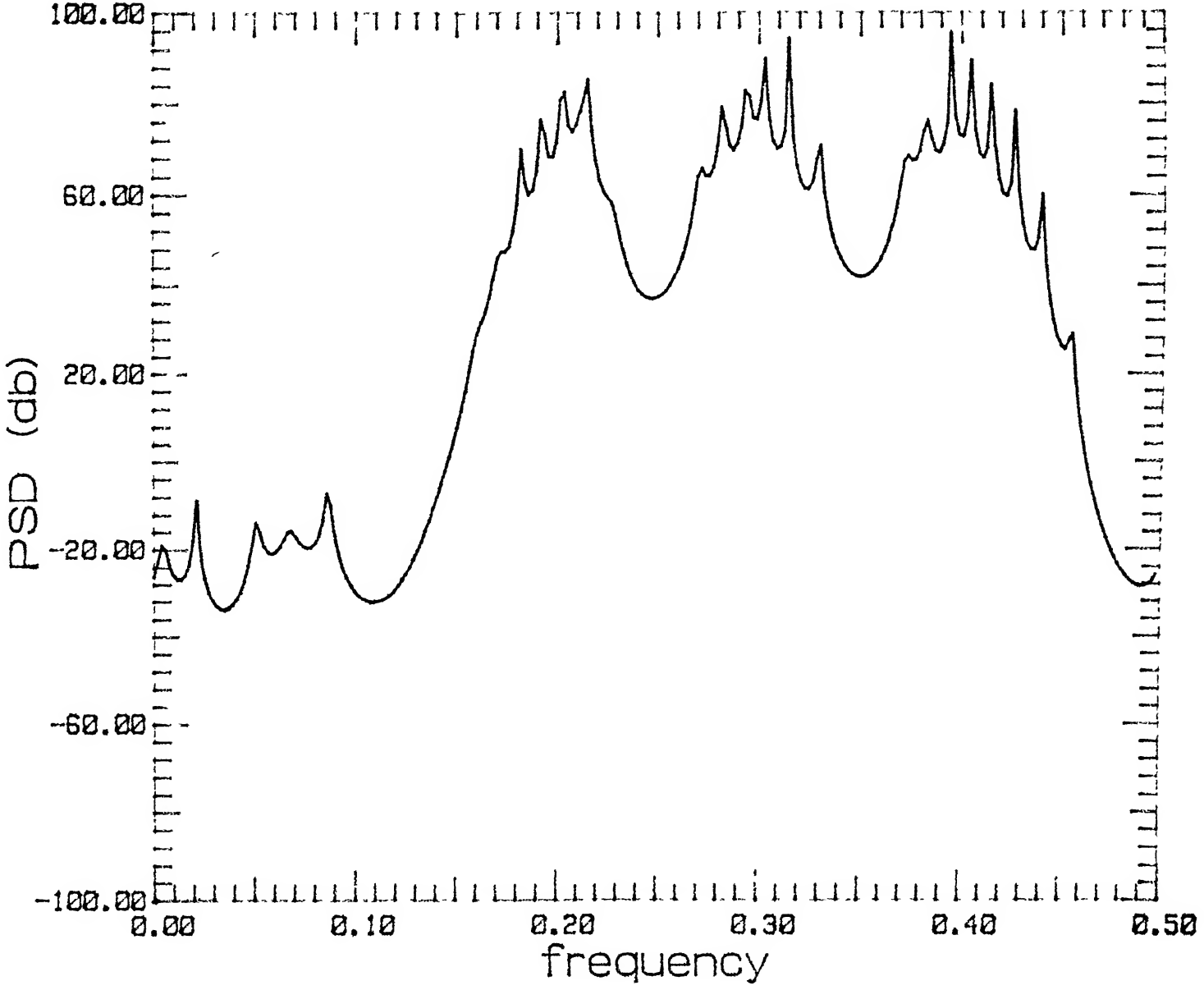


Fig. 3-2-5b PSD : K = + 100, Modal order (2

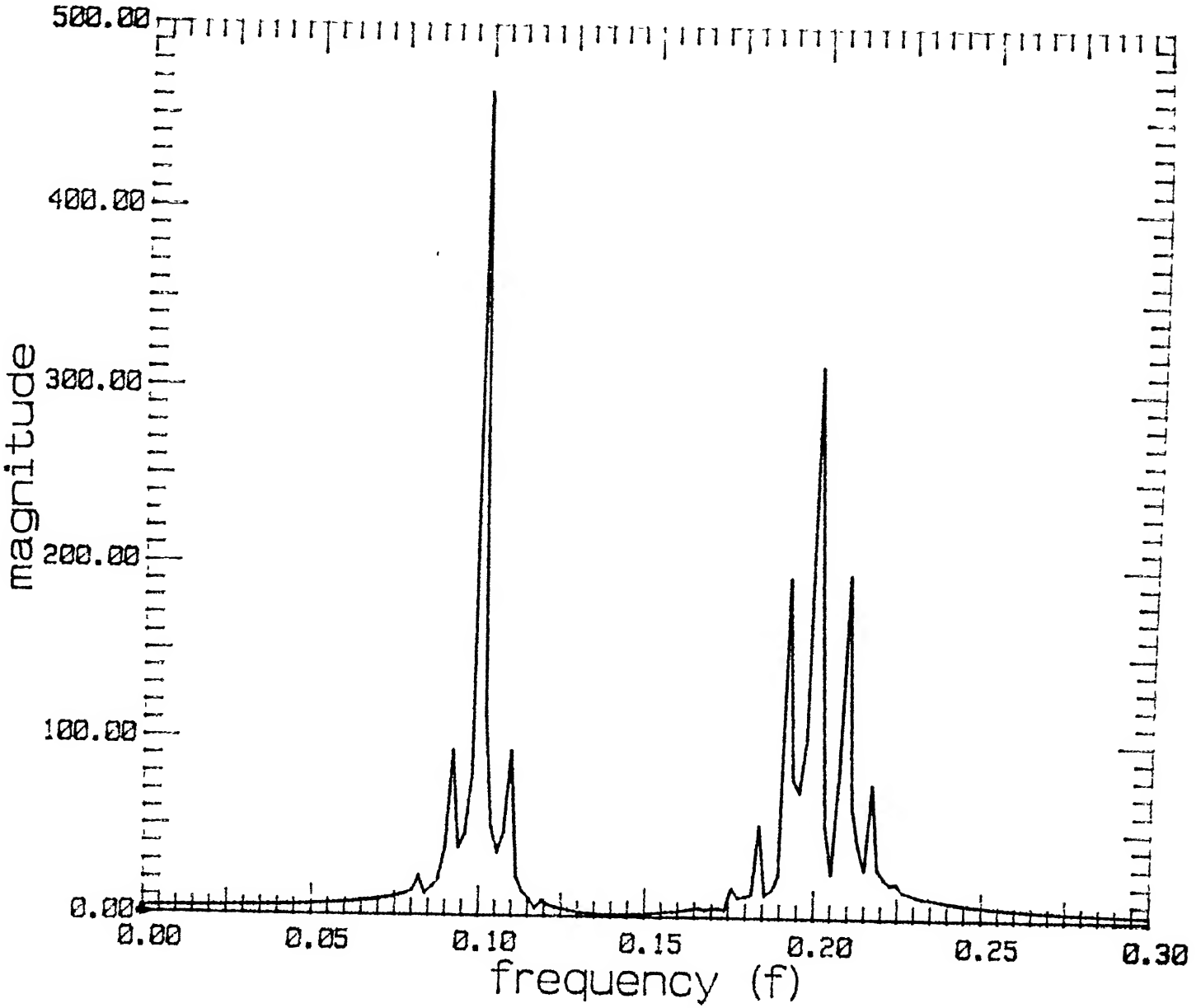
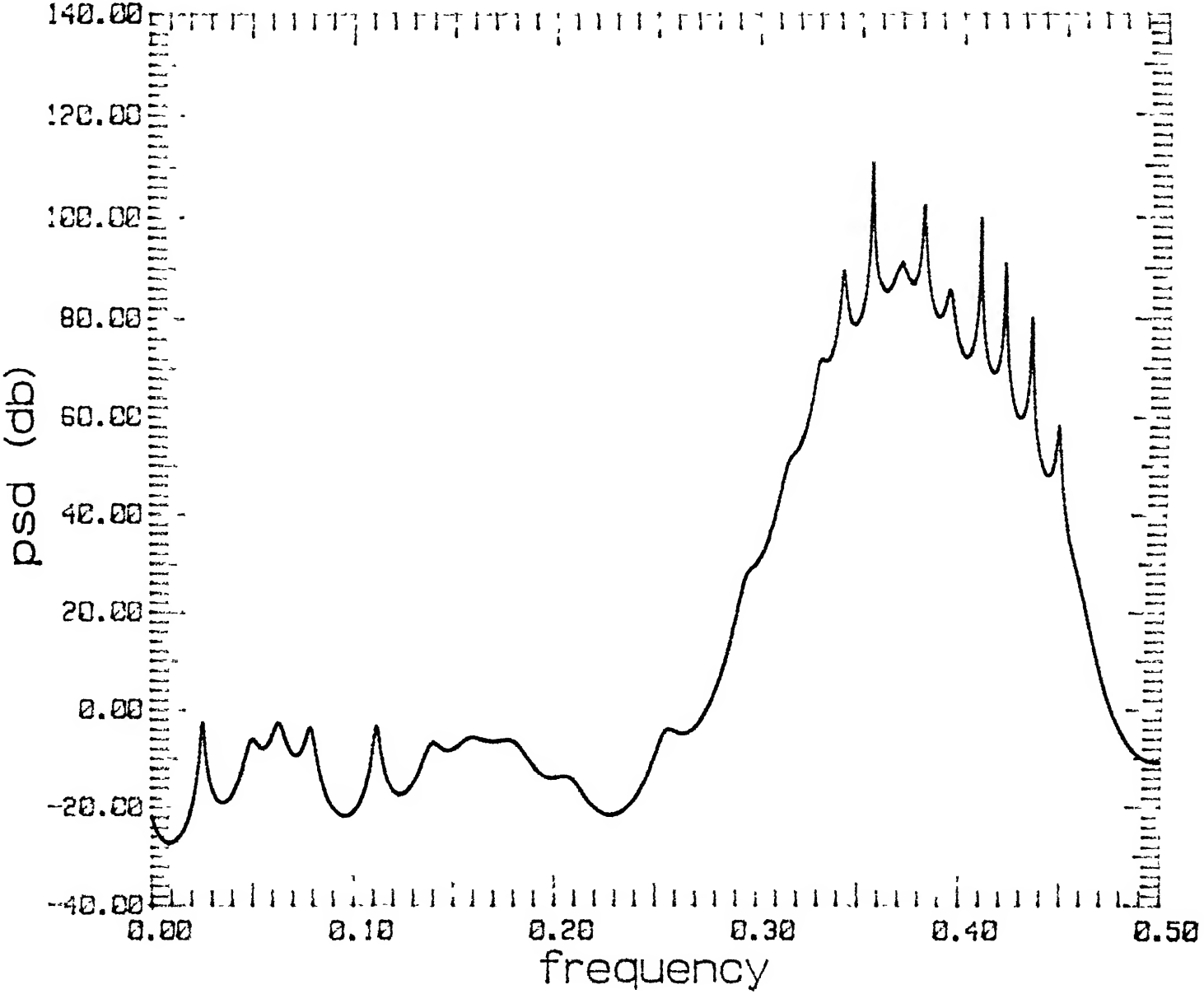


Fig. 3.2.6 Discrete Frequency Transform Magnitude


140.00
120.00
100.00
80.00
60.00
40.00
20.00
0.00
-20.00
-40.00
psd (db)
0.00
0.10
0.20
0.30
0.40
0.50
frequency

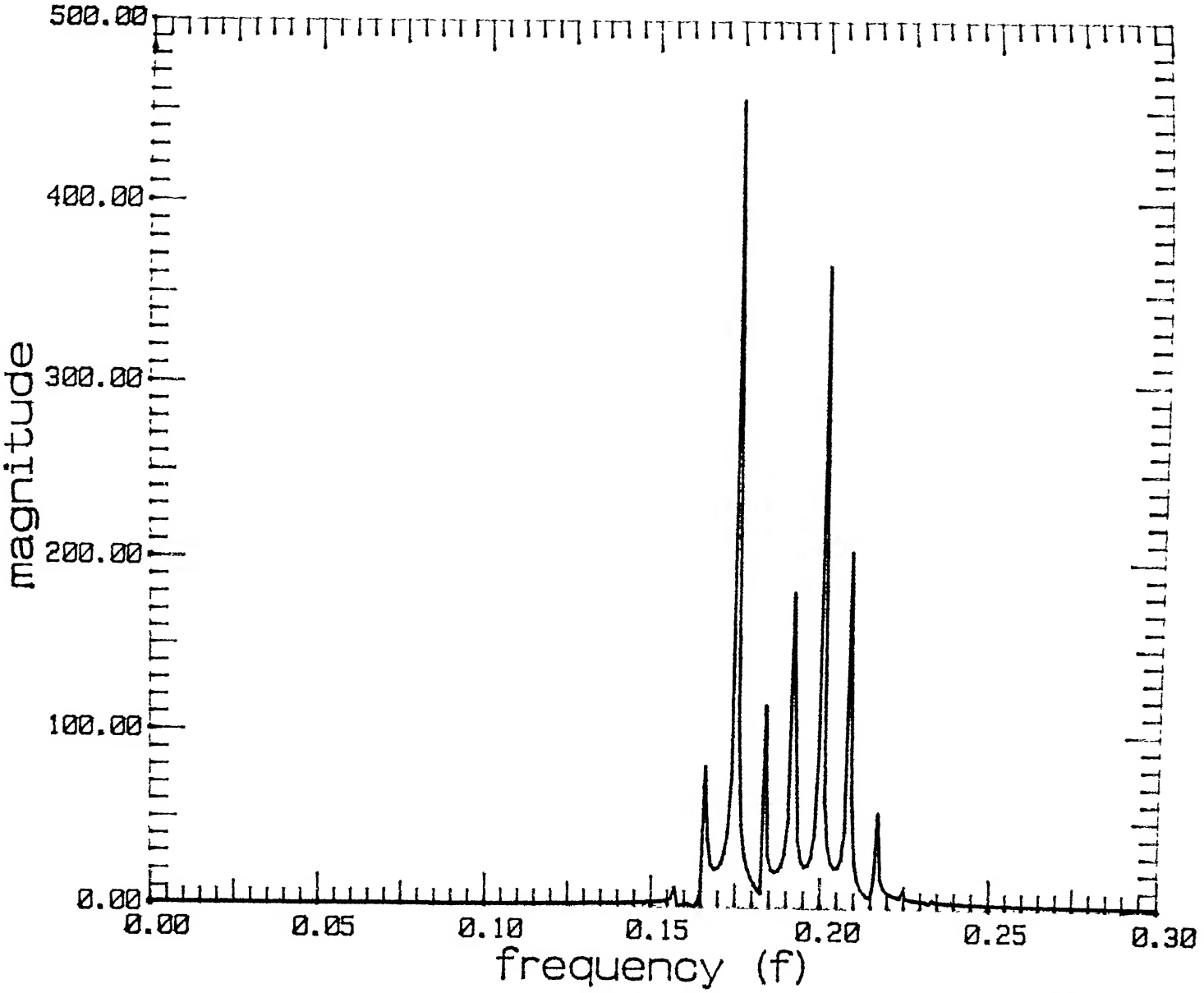


Fig. 3.3.2 Discrete Frequency Transform Magnitude

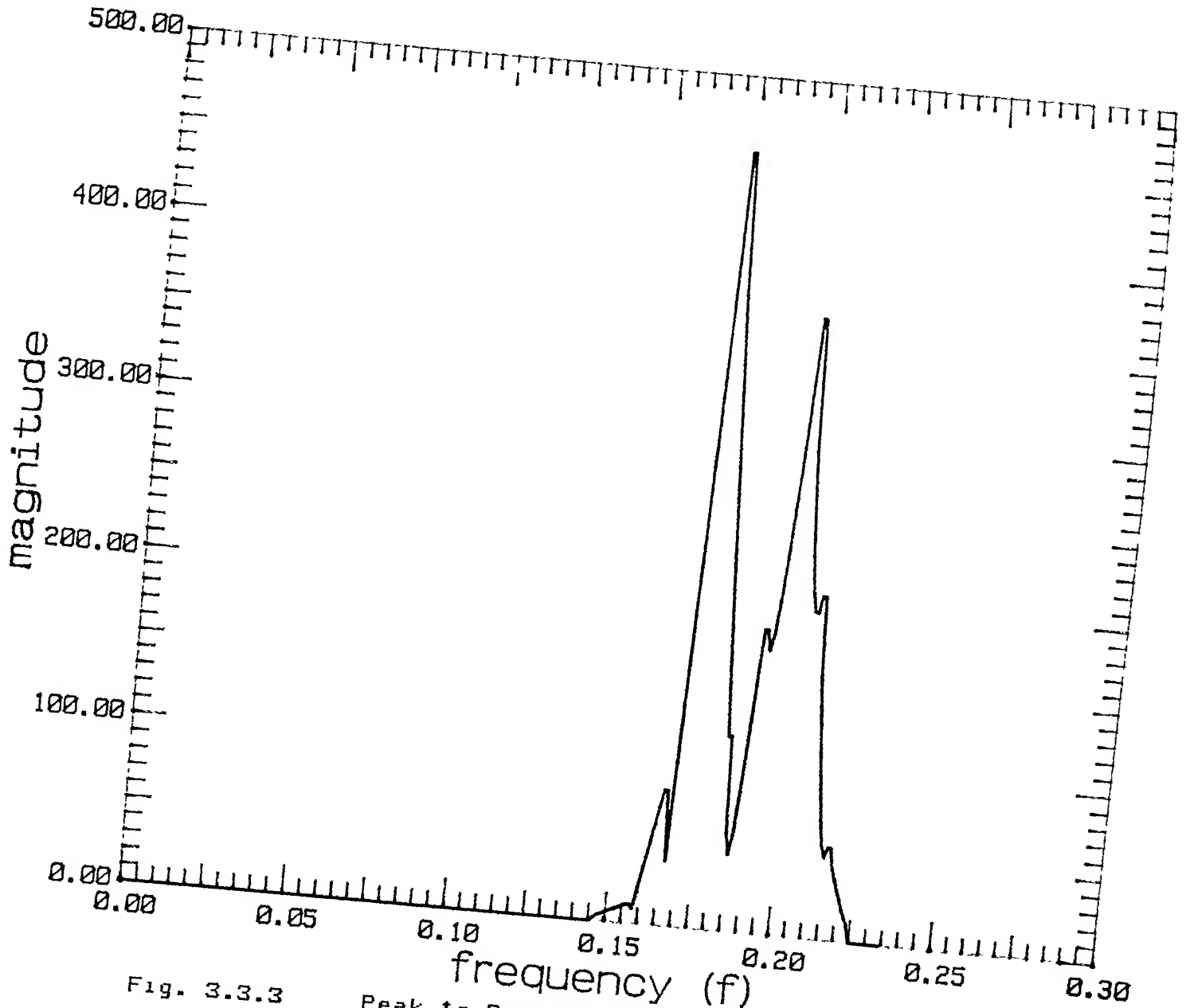


Fig. 3.3.3 Peak to Peak Interpolation

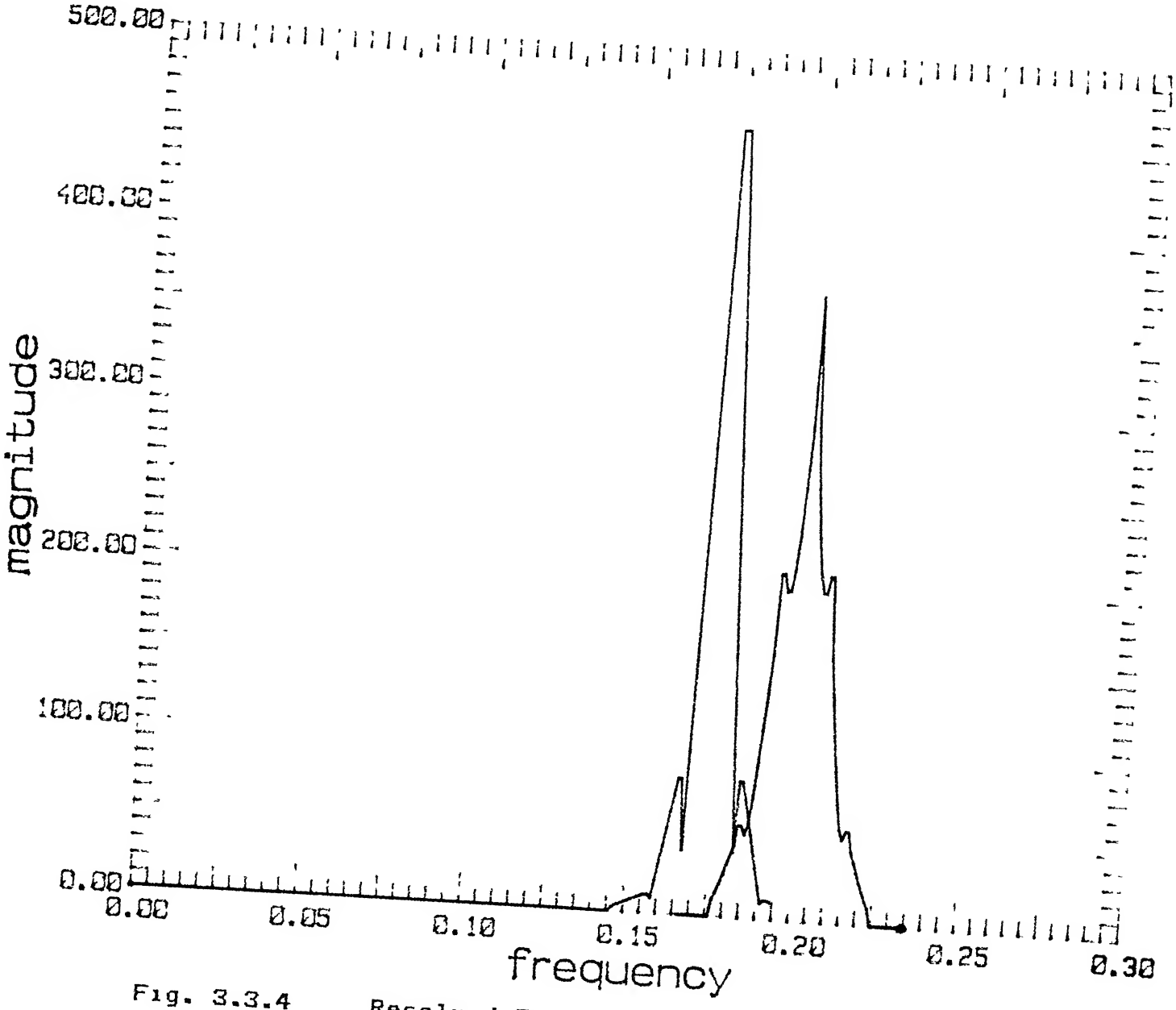


Fig. 3.3.4 Resolved Subsignals

CHAPTER -4

SIMULATION RESULT WITH REAL TIME DATA

4.1 INTRODUCTION

As described in chapter 3 , the parameter estimation is found to be reasonably accurate for the model of sum of complex FM signals. Also the procedure of estimation is shown to be performing well for the overlapping patterns case. The next step will be to try to fit some real time non-stationary data on this model.Speech signal is an example of a non-stationary signal showing frequency variations and hence was considered as a good candidate for this model.For this purpose, the model was fitted on one representative signal from each voiced and unvoiced speech signal.The representative signals were phonemes \k\ and \e\ respectively.

4.2 SIMULATION RESULTS WITH SPEECH DATA:

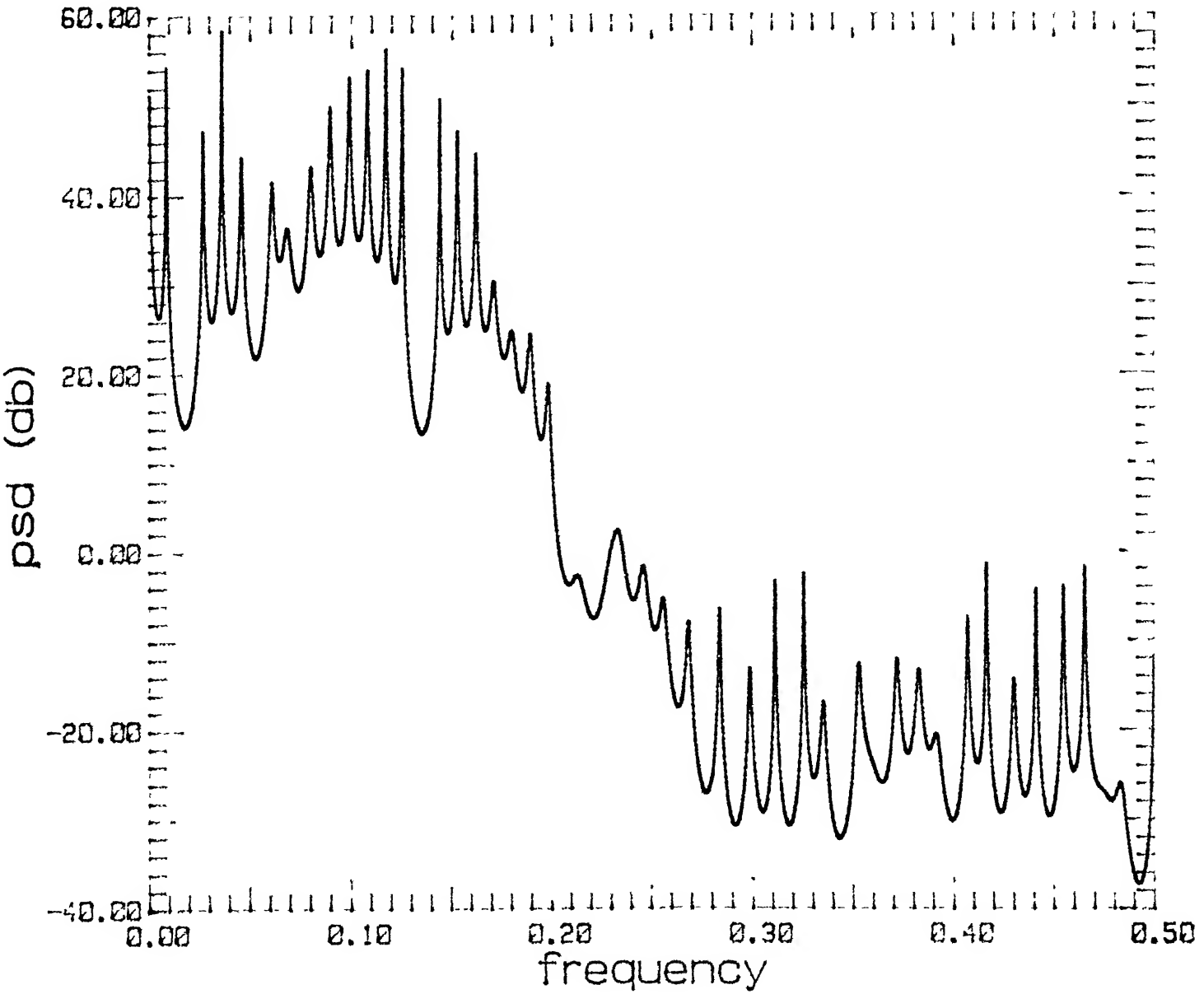
The speech data on which the model was fitted , consisted of unvoiced \k\ and voiced \e\ .A total of 1500 sample points uniformly sampled at 16 KHz for each phonemes were available ie the duration was slightly more than 90 ms.The Product Function was calculated for lag of -200 to + 200 .These were fitted onto an AR Model and the PSD was computed . Starting with a low model order of 10 and then increasing it,the correct model order was established.

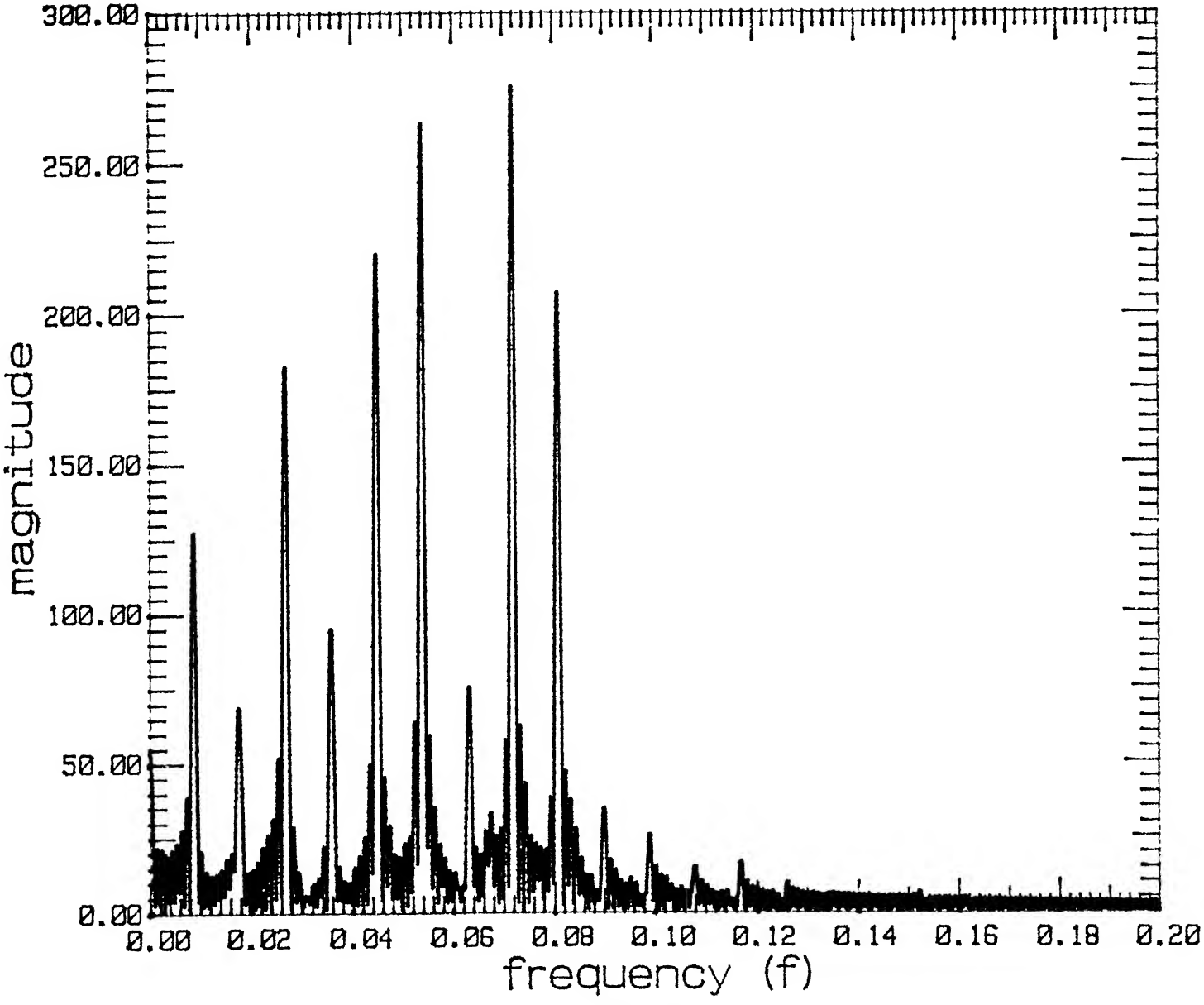
The PSD of unvoiced \k\ sound is shown in fig 4.2.1. Since the data are real valued and so are the Product Functions , the PSD plot is symmetrical about f=0. The DFT magnitude plot is shown in Fig. 4.2.2 . From the PSD , taking aid of the DFT plot , the various FM carrier and modulating frequencies are determined . The spectrum consists of five FM sub-signals . From the DFT plot it is also clear that due to overlaping , many of the peaks have altered amplitudes. Also the peaks about the carriers are not at modulating frequency distance, indicating the poor resolutin of DFT. Dealing the DFT available , by method described in section 2.7 , we find the envelope of the peaks as shown in Fig.4.2.3 . Flipping about the estimated carrier frequency's the pattern of each FM signal is separated. This is shwon in Fig. 4.2.4 .From each pattern, value of corresponding modulatin index is found. Using these frequencies and modulation indexes , we obtain the amplitudes and phases for the constituent FM signals. Using these paramenters we regenerate the consonant sound. Fig 4.2.5 shows the original and regenerated consonant sound.

The PSD of voiced \e\ sound is shown in fig 4.2.6. Since the data are real valued and so are the Product Functions , the PSD plot is symmetrical about f=0 . The DFT magnitude plot is shown in Fig. 4.2.7 . From the PSD , taking aid of the DFT plot , the various FM carrier and modulating frequencies are determined . The spectrum consists of two FM sub-signals . From the DFT plot it is also clear that due to overlaping , many of the peaks have altered amplitudes. Also the peaks about the carriers are at modulating frequency distance . Dealing the DFT available , by method

described in section 2.7 , we find the envelope of the peaks as shown in Fig.4.2.8. Flipping about the estimated carrier frequency's the pattern of each FM signal is separated. This is shwon in Fig. 4.2.9 . From each pattern, value of corresponding modulatin index is found. Using these frequencies and modulation indexes , we obtain the amplitudes and phases for the constituent FM signals. Using these parame-ters we regenerate the consonant sound. Fig 4.2.10 shows the original and regenerated consonant sound.

The model has been shown to be suitable for modelling speech signals with fair degree of success. The model seems to perform better for the unvoiced speech case. This is expected as the modulation in amplitude is much less in unvoiced speech vis-a-vis voiced speech. In fact , by looking at the DFT plot, a fair idea can be obtained about the suitability of a particular signal to be modelled by this particular model. If the DFT plot consists of the desired patterns or approximates it, then the signal is a good candidate for this model. This is amply evident from the simulation cases dealt with.


60.00
40.00
20.00
0.00
-20.00
-40.00
psd (db)
0.00
0.10
0.20
0.30
0.40
0.50
frequency



Discrete Frequency Transform /K/

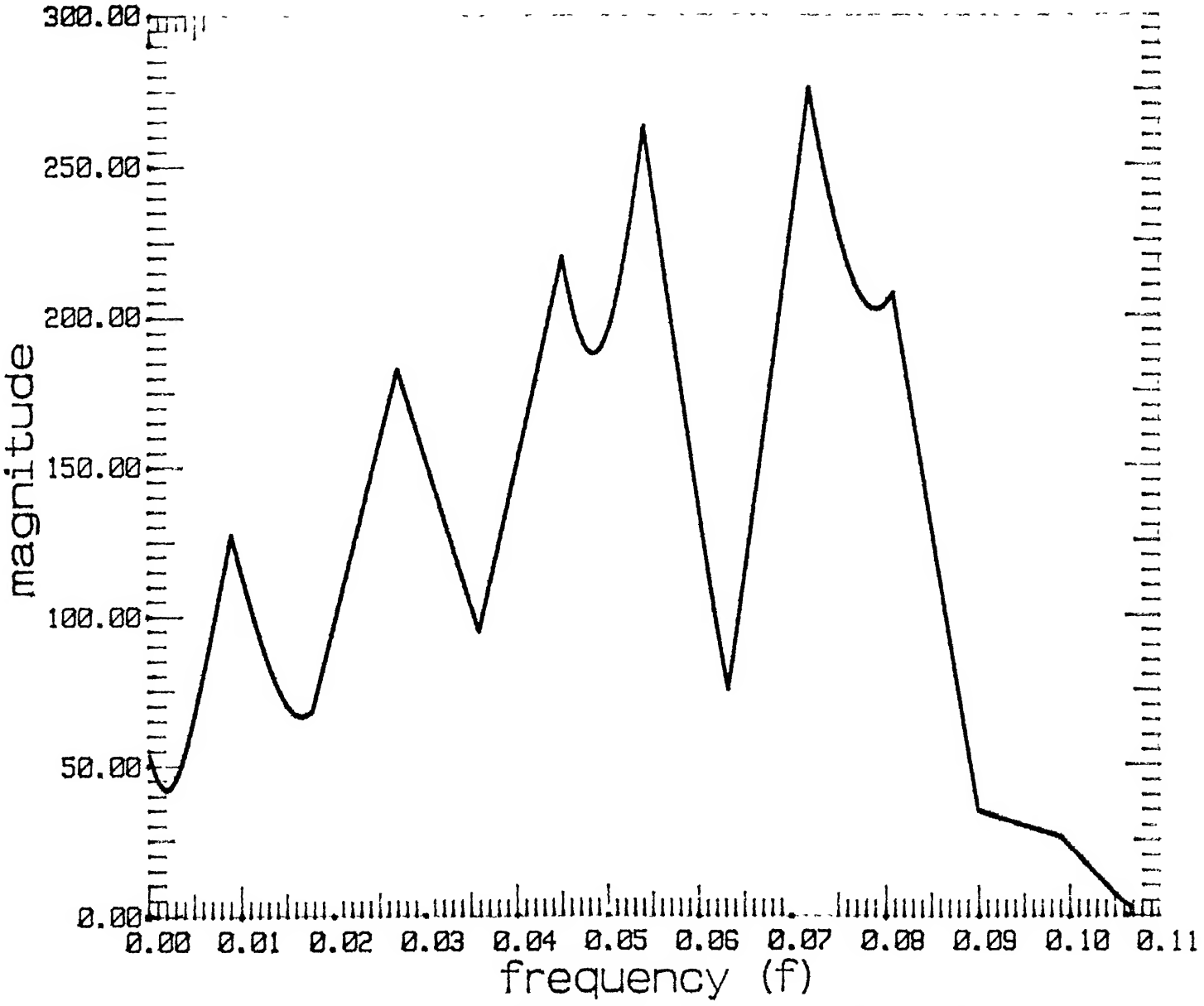
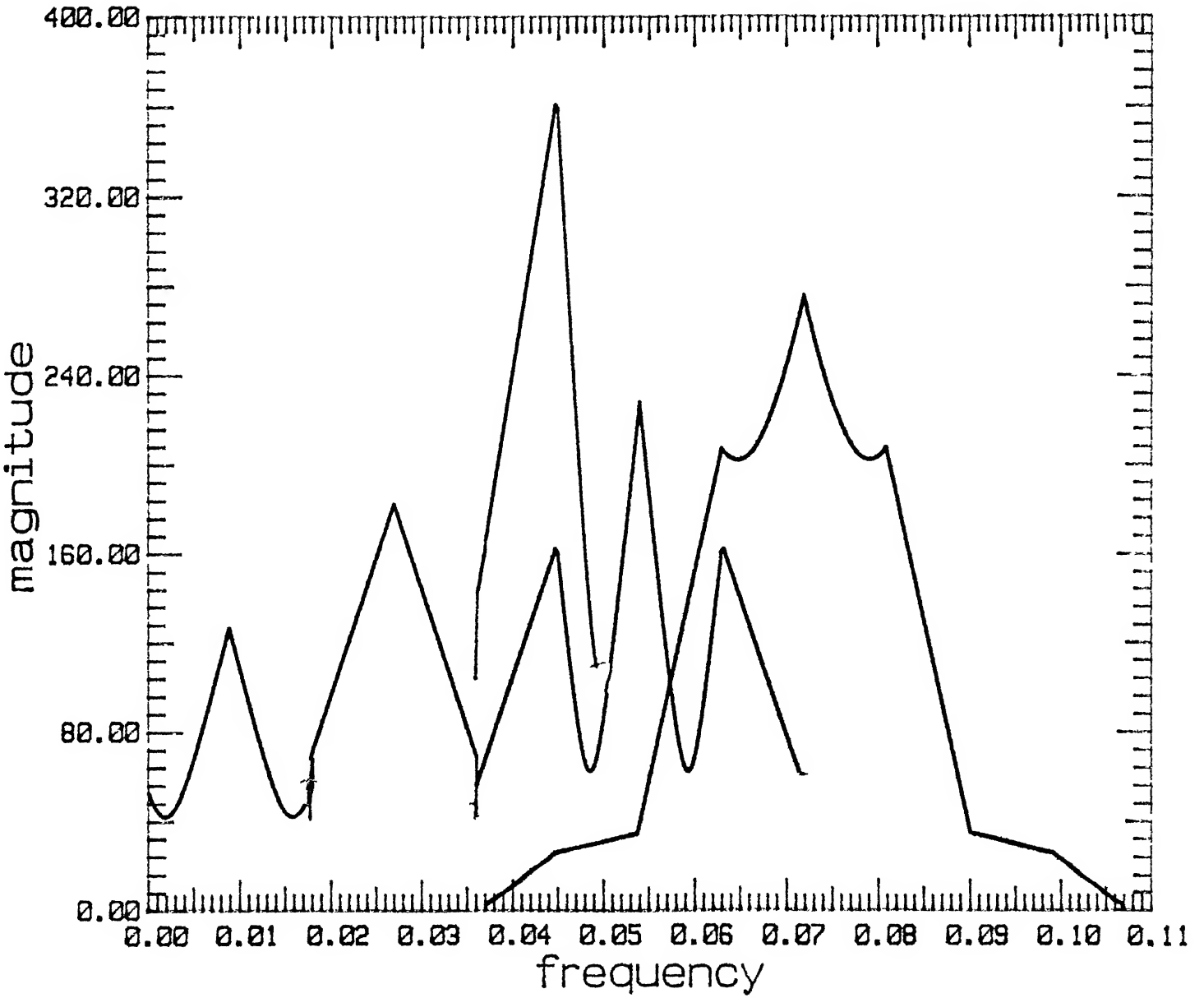


Fig. 4.2.3 Peak to Peak Interpolation /K/



Resolved Subsingals /K/

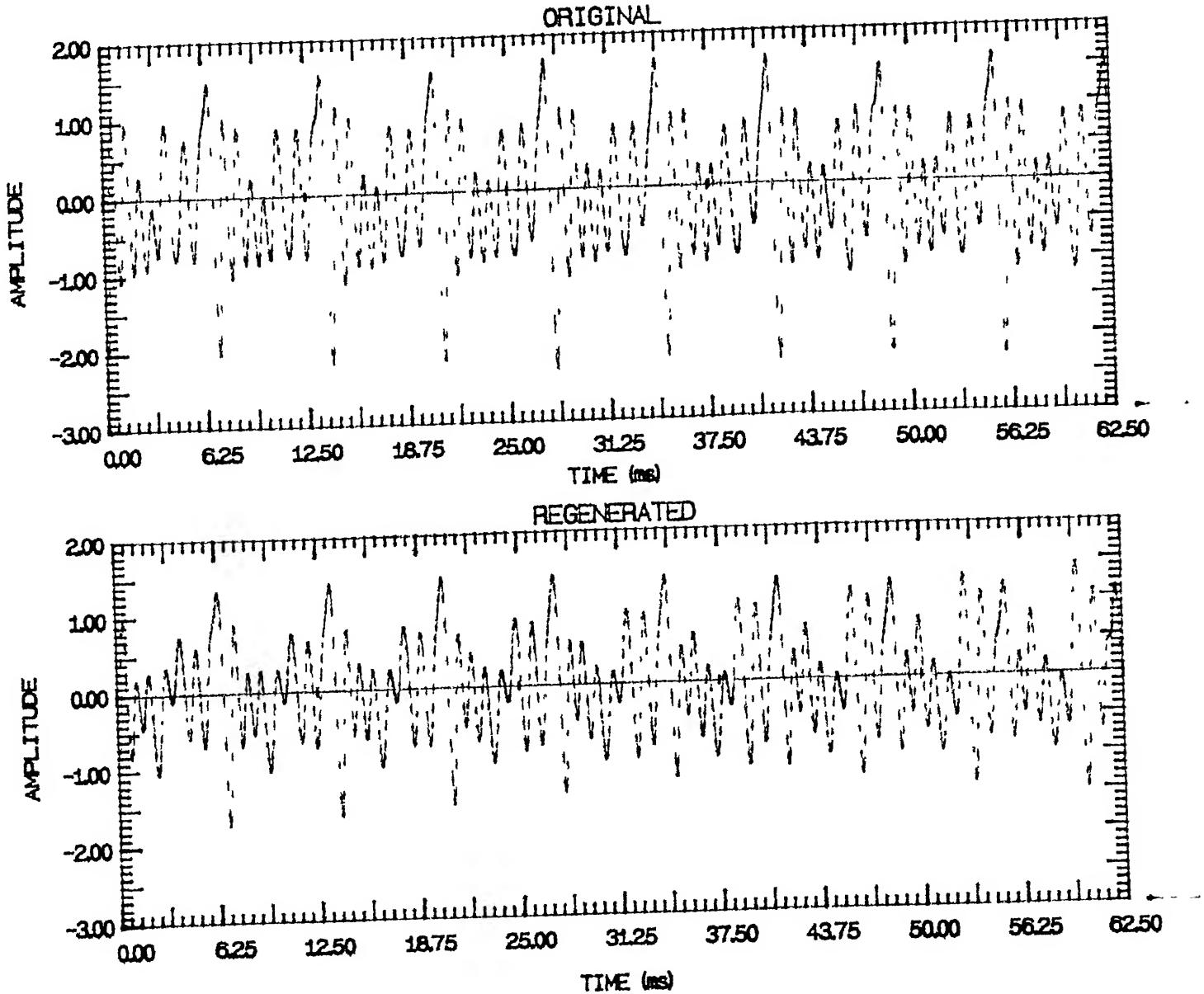


Fig. 4.2.5 Original and Regenerated Phoneme /K/

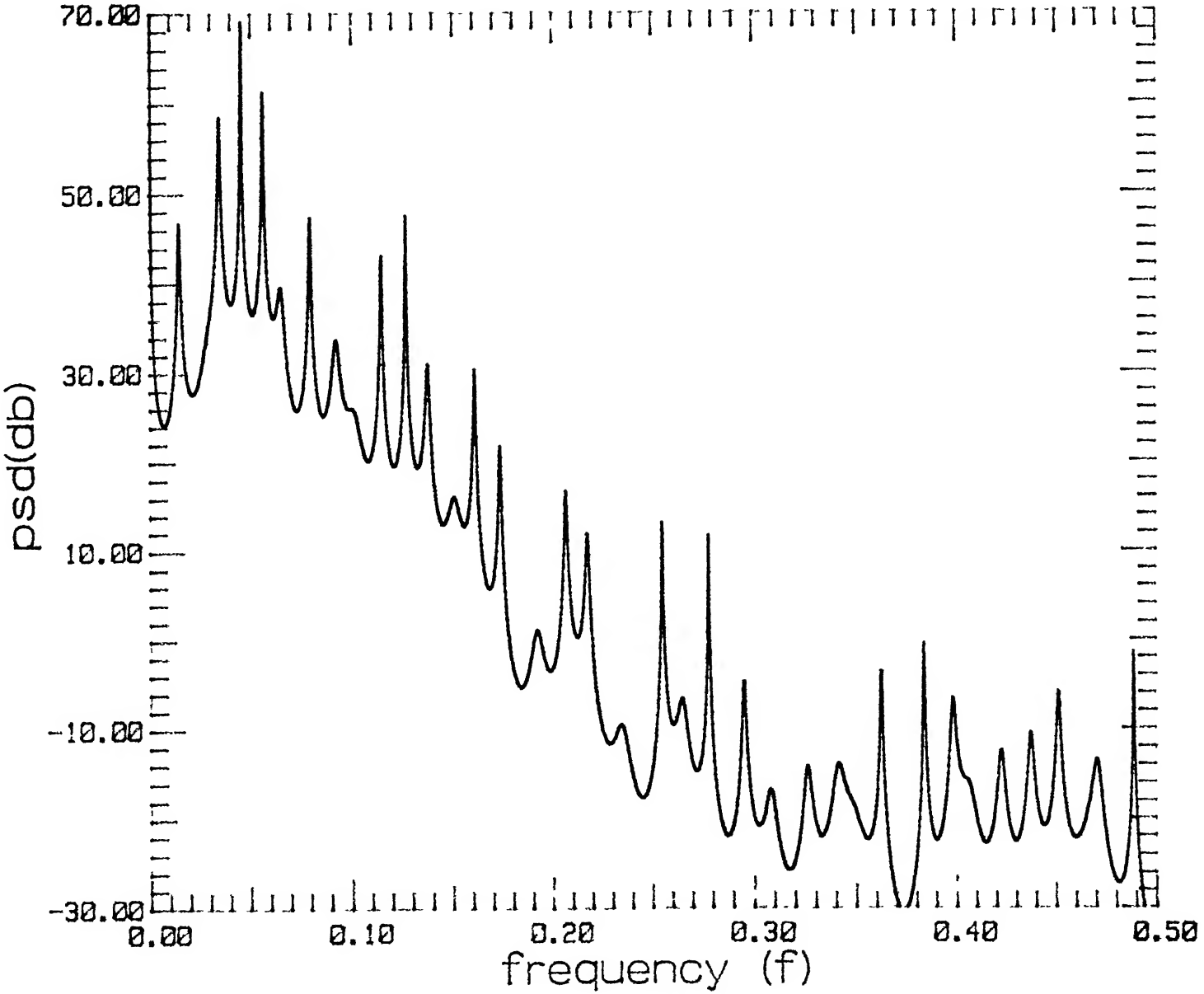


Fig. 4.2.6 PSD Phoneme / e/

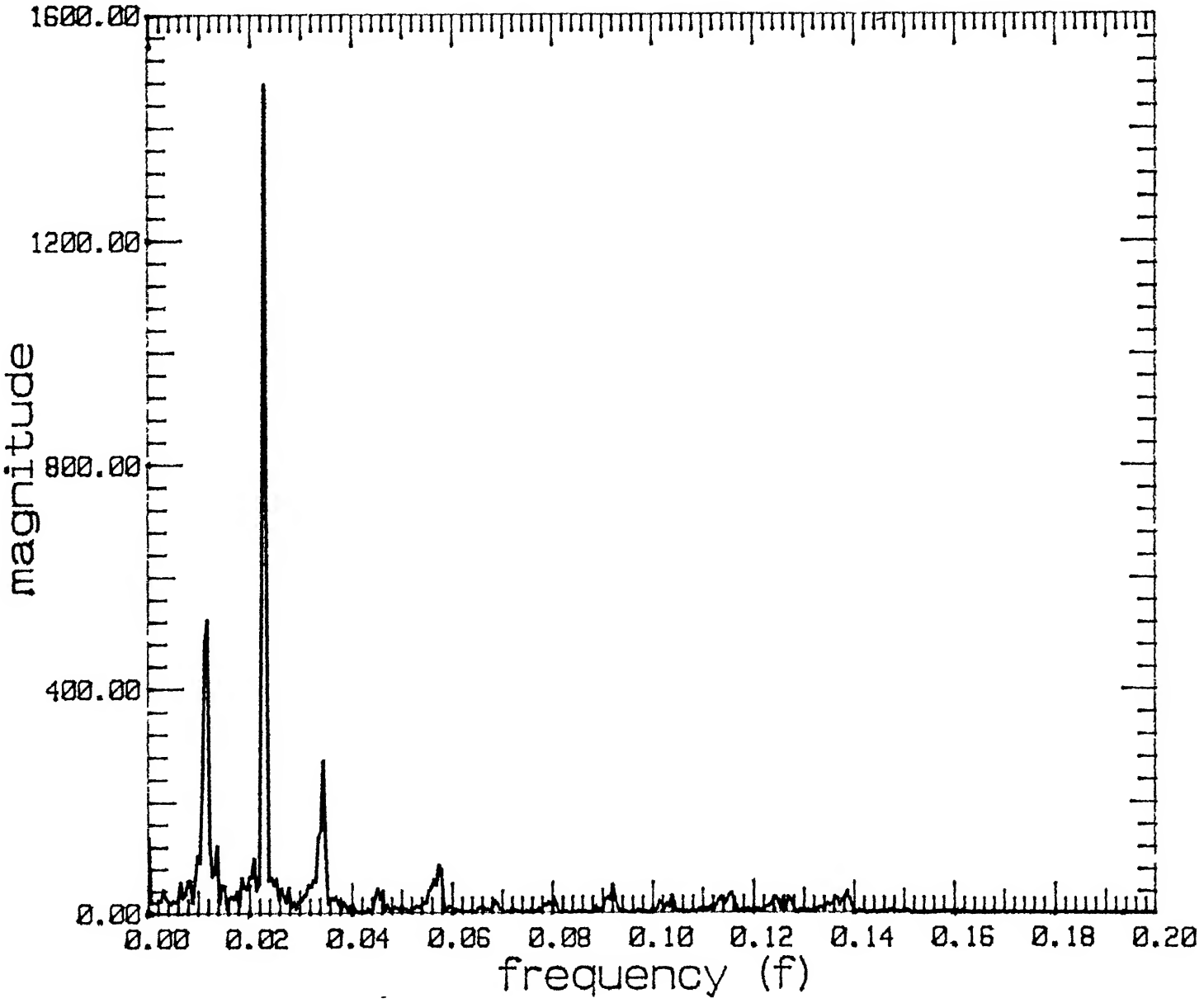


Fig. 4.2.7 Discrete Frequency Transform /e/

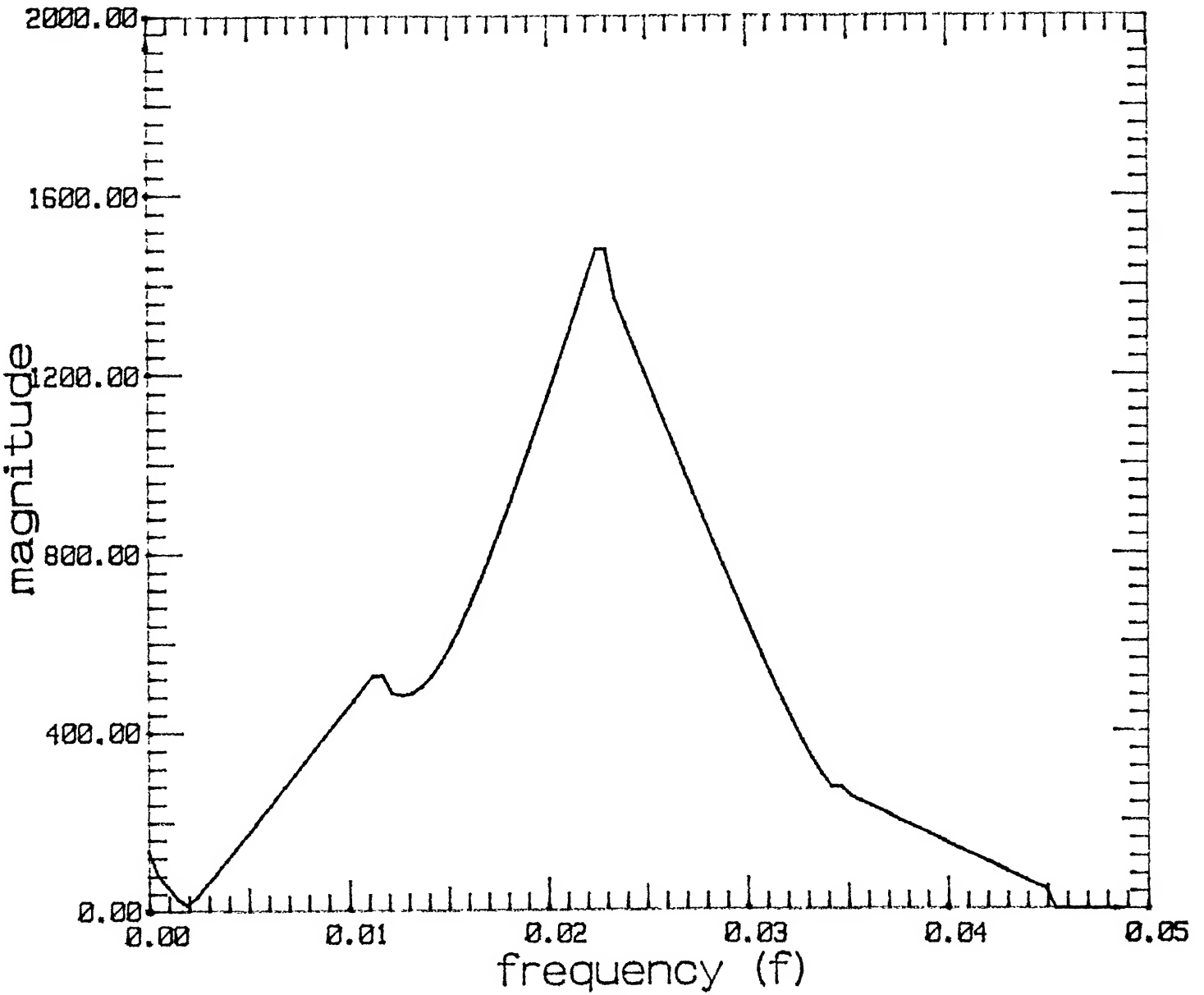


Fig. 4.2.8 Peak to Peak Interpolation /e/

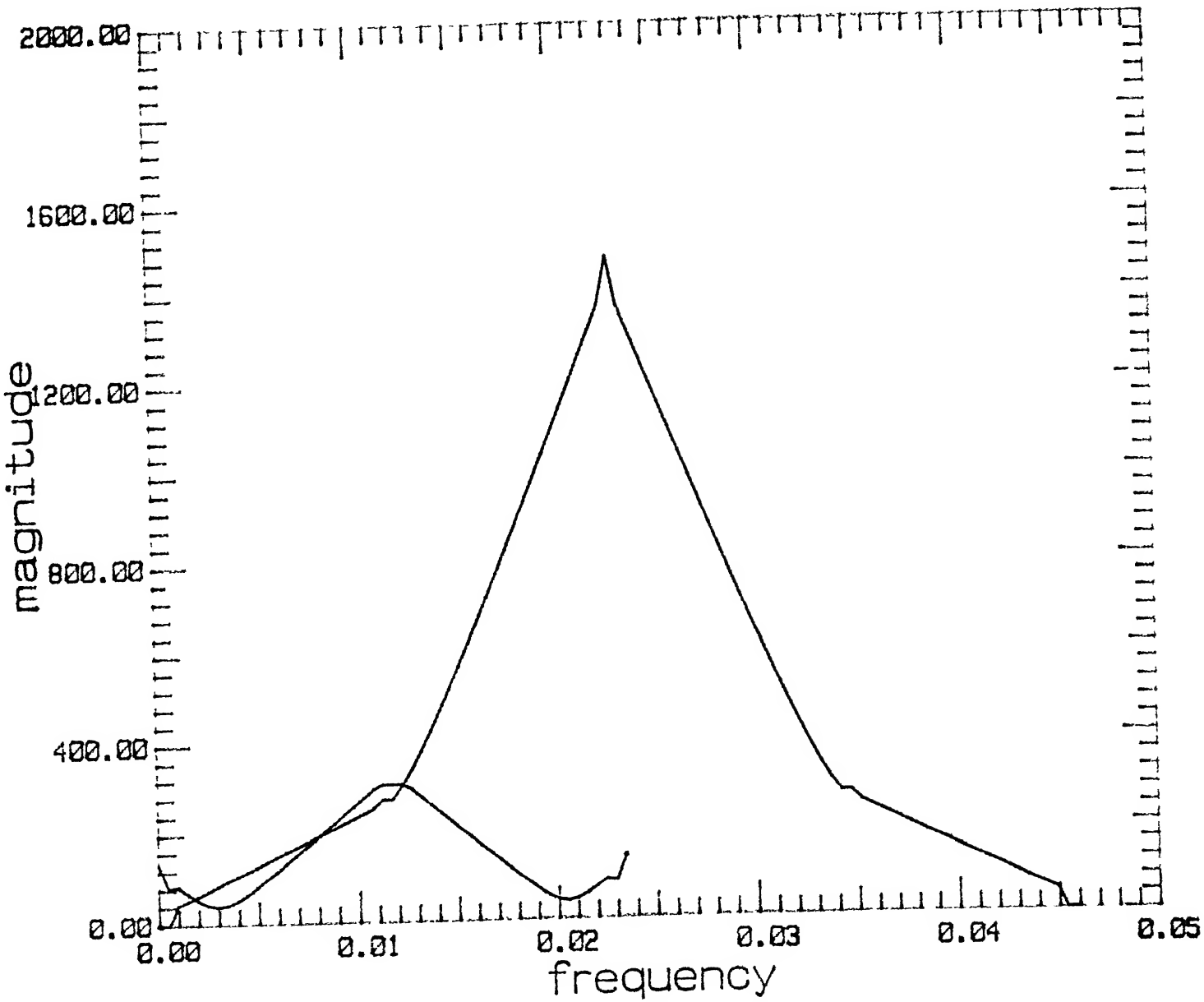


Fig. 4.2.9 Resolved Subsingals /e/

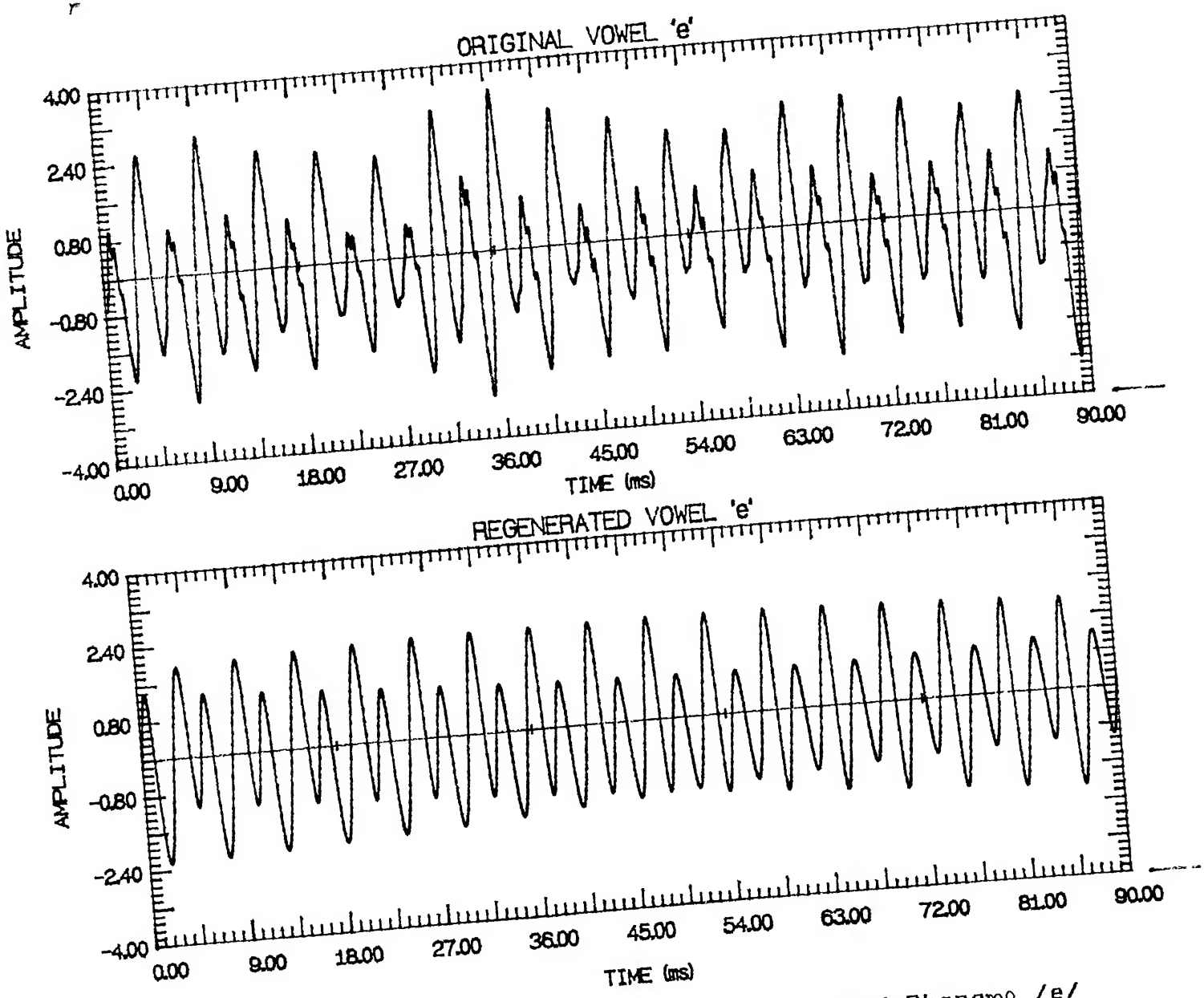


Fig. 4.2.10 Original and Regenerated Phoneme /e/

CHAPTER-5

CONCLUSION

The model introduced in this thesis can be used to model signal data from a deterministic or a stochastic process , as a sum of several FM signals.With synthesized data, reasonably accurate estimation of parameters are obtained.

The Fourier transform based techniques though limited in resolution, perform adequately under low SNR. AR modelling has the advantage of high resolution spectral estimation, though its performance is good at high SNR and deteriorate at low SNR .A combination of the two, using the capability of each makes the model parameter estimation feasible accurately.

The model has been found particularly suitable for speech signals. Knowing the parameters of a signal the resulting parameters can be used as a feature vector for pattern recognition. The application of this model may include speech segmentation, speaker verification, speaker identification, speech recognition and language identification etc.

However a much more detailed experimentation is required to fully exploit the capabilites of this model. Specifically a more general model consisting of a sum of AM-FM signals may be suitable to model any type of non-stationary signal. However, more work

will be needed in this direction before the method can be standardised.

REFERENCES

1. A. Papoulis, "Probability, Random Variables and Stochastic Processes", Mc Graw-Hill, New York, 1984.
2. S. Haykin, "Adaptive filter Theory", Prentice Hall, Englewood Cliffs, N J, 1986.
3. S. M. Kay, "Modern Spectral Estimation: Theory and Application", Prentice Hall, Englewood Cliffs, N. J., 1988.
4. S. M. Kay and S. L. Maple, Jr., "Spectrum analysis - A Modern Perspective, Proce IEEE, Vol. 69, pp 1380-1419, Nov, 1981.
5. Proakis John G, Manolakis Dimitris G., "Digital Signal Processing-Principles, Algorithms and Applications", Maxwell Macmillan International Edition, 1992.
6. Elie. J. Baghdady ed., Lectures on Communication system theory, McGraw-Hill, New York.
7. Ed. F. Derettere ed., SVD and Signal Processing, North Holland 1988.